5 reasons why you should participate in proficiency tests

Dr. Andreas Balster, Deutsches Institut für Ringversuche

Facts & Figures

► The Kunststoff-Institut Lüdenscheid supports you with
  ▪ the selection
  ▪ the development
  ▪ optimization and implementation
  of products, tools and processes in the entire field of plastics technology

► The institute is financed exclusively by services in the form of consulting, collaborative and development projects

► A group of more than 300 active members from Europe represents the majority shareholder, called “Trägergesellschaft”
Fast, competent solutions for the plastics industry

2002: First round robin tests of the Kunststoff-Institut, originating from our own needs
2006: First participant outside Germany
2010: First participants outside Europe
2016: For the first time over 400 participating laboratories
2017:
- Application for the trademark "Deutsches Institut für Ringversuche" (German Institute for Proficiency testing)
- Application for accreditation according to DIN ISO/IEC 17043
WHAT ARE PROFICIENCY TESTS?

► Proficiency tests represent a possibility of external quality assurance
► A group of laboratories is assigned a measurement, testing or analysis task
► Each participating laboratory receives
  ▪ the same sample
  ▪ same information
  ▪ the same period of time for implementation

What are Proficiency tests?

Deutsches Institut für Ringversuche
Deutsches Institut für Ringversuche

A Proficiency test is organised and administered from a central location. This one cares…

- to select and describe and organise the respective procedure in a technically correct manner
- to select, prepare and ship suitable sample materials
- registration and supervision of the participating laboratories, in particular with regard to the protection of the environment
  - neutrality and objectivity
  - the anonymity of the participants
- in order to ensure that data transmission is as error-free as possible
- for a correct statistical evaluation of the results
- to ensure meaningful data preparation and interpretation

What are Proficiency tests?

An interlaboratory test is usually carried out according to the following scheme:
How is the result of interlaboratory tests reported?

- Each procedure is subject to a certain possible attainable level of precision, which can be seen from the totality of the results of all participants.

- After elimination of outliers, the **best estimate of the true value** is determined by averaging.

- Each laboratory shows in its results a certain deviation from the *best estimate*, the extent of which expresses the laboratory's performance.

- The result is given in such a way that the error caused by the method and the material samples is excluded.

WHY PROFICIENCY TESTS?
Reason 1:

For laboratories accredited according to DIN ISO/IEC 17025, regular participation in proficiency tests is obligatory.

DIN EN ISO/IEC 17025

7.7.2

The laboratory shall monitor the quality of the laboratory performance by comparing with output of other laboratories, where available and appropriate. This monitoring shall be planned and reviewed and shall include, but not be limited to a selection from the following list:

a) participation in proficiency testing;

[...]

b) participation in interlaboratory comparisons other than proficiency testing.
Reason 2:

Results from round robin tests can be used (e.g. for customers) as proof of competence

Ringversuche als Marketinginstrument

► An accurate laboratory performance confirmed by an independent party is the best advertisement a commercial laboratory can wish for.

► In safety-critical applications and industries, inadequate laboratory performance can be an exclusion criterion.
  ▪ This is not an argument against proficiency testing 😊

► The mere fact that a laboratory voluntarily undergoes a performance review may be decisive for a client

Quelle: Kunststoff-Institut Lüdenscheid
Reason 3:

There’s no better way to objectively check your own laboratory performance

► How do you know that what you are measuring is correct?
  ▪ Factory calibration?
  ▪ DAkkS calibration certificates?
  ▪ These only indicate that a device has been calibrated.
  ▪ Calibration is usually performed on traceable standards, i.e. on samples with known results.
  ▪ (external) calibration is carried out by persons who have been commissioned with it
  ▪ Calibration is carried out at regular intervals, usually internally more often than externally
  ▪ Calibration checks the instrument - no more and no less
How do you calibrate your system?

- Reference materials can be used at any time to test a process if:
  - they are available in sufficient quantity
  - they remain stable over a longer period of time, and
  - their subsets are homogeneous, i.e. deliver the same results

For non-destructive tests, calibration can be repeated as often as required.

Standards provide a known result (e.g. melting point) in a predictive interval.

Standards

- A standard usually provides only one "breakpoint" along the entire possible measuring range
- Danger of a "bias" of the performer if the result is known
- Loss, destruction, alteration or contamination of a standard can lead to massive problems
- Many calibrations run in a special mode of the system; the real case is not covered by it
What is different about proficiency testing?

- Interlaboratory tests are carried out under regular laboratory conditions
- Samples that are distributed for the performance of the measurement task (industry relevant) are checked and monitored with regard to their suitability
- For destructive tests, the samples are sufficiently homogeneous, i.e. lead to the same results
- The material is stable over the period of the tests and does not change its properties
- The interval of possible results is variable and not generally known to the operator

Reason 4:

Proficiency tests provide insight into the performance of processes and equipment
Example: DSC Analysis (ISO 11357-2)

- **Determination of $T_g$** (Inflection point of the curve)
- **Length of arrows:** Deviation from assigned value (in units of standard deviation - $z_u$-scores)
  - Blue: acceptable ($|z_u| \leq 2$)
  - Yellow: questionable ($2 < |z_u| < 3$)
  - Red: unsatisfactory ($|z_u| \geq 3$)
- **Direction of arrows:** Direction of deviation

### Table: Laboratory Mean Values

<table>
<thead>
<tr>
<th>Lab</th>
<th>PA PACM</th>
<th>PMMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>108</td>
<td>124,05 °C</td>
<td>112,39 °C</td>
</tr>
<tr>
<td>109</td>
<td>136,31 °C</td>
<td>114,08 °C</td>
</tr>
</tbody>
</table>

- Participants 108 and 109 come from the same laboratory
- There are two different devices and the same operator
- It can be known that at least one result must be wrong – but not necessarily which one *
- After the evaluation, it is clear that 108 systematically delivers values that are too low
- Analogous to this constellation, different operators can also be tested
- If further reference material is purchased, investment decisions can also be simplified at a later date

*: In this case, the approximate correct result can be found out.
Reason 5:

Proficiency tests are an ideal opportunity to train the personnel

What can go wrong?

- **Rough mistakes**
  - can always be avoided;
  - Occurs e.g. due to the use of damaged measuring instruments or incorrect reading

- **Systematic faults**
  - Don’t just happen by accident, but with "system"
  - Measured values are biased in the same way
  - With an incorrectly calibrated measuring instrument, all values are either too high or too low
  - Systematic errors can be avoided

- **Random errors**
  - In contrast to the gross and systematic mistakes, these can never be eliminated completely
  - They are created by a multitude of different, uncontrollable actions and are always distributed without rules
What am I doing wrong?

- At least two measurements shall be taken to determine whether an error occurs randomly or systematically
  - All our proficiency tests include the examination of (at least) two levels
  - The evaluation in the final report provides information about a possible trend

**Example**: RV-Series 2017, melt index determination acc. to ISO 1133

- 3 levels / samples: Polycarbonate, PE-LD, PP-GF
- Participation of 25 laboratories
- Not every laboratory has examined all three samples
- MVR data is supplied by the device; for MFR, an additional while weighing of the strands is carried out

---

**Example: MVR/MFR (ISO 1133)**

<table>
<thead>
<tr>
<th>Laboratory Code</th>
<th>Melt Flow Rate (g/10min)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>14.3 15.6 16.9 17.2 18.3 19.5 20.7 21.9 23.2 24.5 25.8 27.1 28.4 29.7 31.0 32.3 33.6 34.9 36.2</td>
</tr>
</tbody>
</table>

**Table**:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of data sets</td>
<td>14</td>
</tr>
<tr>
<td>Assigned value</td>
<td>15.43 g/10 min</td>
</tr>
<tr>
<td>Expanded uncertainty</td>
<td>0.64 g/10 min</td>
</tr>
<tr>
<td>Standard deviation for proficiency assessment</td>
<td>1.26 g/10 min</td>
</tr>
</tbody>
</table>

---

**Diagram**

- Lab mean and lab standard deviation
- Mittelwerte der einzelnen Prüflisten
- Test series-specific mean values
- Assigned value and extended uncertainty
- Score = +2
Example: MVR/MFR (ISO 1133)

**PE-LD (natural) | PE-LD (natural)**

- Anzahl Datenpunkte $p$ | Number of data points $p$
- Zumindestens Wert $x_p$ | Assigned value $x_p$
- Erweiterte Unsicherheit des zumindesten Wertes $u(x_p)$ | Extended uncertainty of assigned value $u(x_p)$
- Standardabweichung für die Differenzansicht $s_p$ | Standard deviation for proficiency assessment $s_p$

**PP-GF (natural) | PP-GF (natural)**

- Anzahl Datenpunkte $p$ | Number of data points $p$
- Zumindestens Wert $x_p$ | Assigned value $x_p$
- Erweiterte Unsicherheit des zumindesten Wertes $u(x_p)$ | Extended uncertainty of assigned value $u(x_p)$
- Standardabweichung für die Differenzansicht $s_p$ | Standard deviation for proficiency assessment $s_p$
The laboratories
- 027
- 041
- 115
- 131 and
- 139
show different results in MFR determination

Overview of $z_0$-scores for MFR measurement

<table>
<thead>
<tr>
<th>Lab</th>
<th>PC</th>
<th>PE-LD</th>
<th>PP-GF</th>
<th>Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>027</td>
<td>-2,81</td>
<td>-0,43</td>
<td>1,99</td>
<td>n. syst.</td>
</tr>
<tr>
<td>041</td>
<td>0,76</td>
<td>-0,24</td>
<td>121,30</td>
<td>n. syst.</td>
</tr>
<tr>
<td>115</td>
<td>0,59</td>
<td>0,89</td>
<td>0,42</td>
<td>syst.</td>
</tr>
<tr>
<td>131</td>
<td>-0,66</td>
<td>-1,08</td>
<td>-0,58</td>
<td>syst.</td>
</tr>
<tr>
<td>139</td>
<td>-0,29</td>
<td>0,42</td>
<td>-0,22</td>
<td>n. syst.</td>
</tr>
</tbody>
</table>

Conclusion?
- Laboratory 027 does not master the process well in principle
- Laboratory 041 has produced an apparent outlier.
  Recommendation: check results for plausibility
- Laboratory 115: tendency towards high results. Process should be checked for possible influences (temperature, balance, nozzle...)
- Laboratory 131: tendency to poor results. Process should be checked for possible influences (temperature, balance, nozzle...)

Example: MFR/MVR (ISO 1133)

Overview of $z_u$-scores for MVR measurement

<table>
<thead>
<tr>
<th>Labor</th>
<th>PC</th>
<th>PE-LD</th>
<th>PP-GF</th>
<th>Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>027</td>
<td>0,55</td>
<td>2,27</td>
<td>0,83</td>
<td>syst.</td>
</tr>
<tr>
<td>041</td>
<td>1,89</td>
<td>0,25</td>
<td>143,27 (syst.)</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>0,13</td>
<td>0,70</td>
<td>0,40</td>
<td>syst.</td>
</tr>
<tr>
<td>131</td>
<td>-1,50</td>
<td>-1,29</td>
<td>-0,54</td>
<td>syst.</td>
</tr>
<tr>
<td>139</td>
<td>-0,20</td>
<td>0,46</td>
<td>-0,47</td>
<td>n. syst.</td>
</tr>
</tbody>
</table>

Red: Sign change compared to MFR

Conclusion?
- Laboratory 027 does not master the procedure well in principle; both the measuring and weighing processes are affected.
- Systematic deviation, weighing process significantly unsystematic.
- Laboratory 041: Instrument appears to systematically deviate upwards (outlier not taken into account) - Weighing process systematically reduces result.
- Laboratory 115: Weighing procedure has only a very small, yet amplifying influence on systematic deviation.
- Laboratory 131: Weighing procedure has a significant influence on systematic deviation, but this part of the error appears less systematic.

Proficiency tests are a versatile and effective means of quality assurance.
- The DIR offers more than 200 experiments in the plastics sector.
- If a test is missing - please contact us!
- Contact:

www.proficiency-test.info
www.dir-kimw.de/en
Vielen Dank!