USE AND INTERPRETATION OF THE UNCERTAINTIES OF THE PARTICIPANTS IN THE FRAMEWORK OF AN INTERLABORATORY COMPARISON TEST

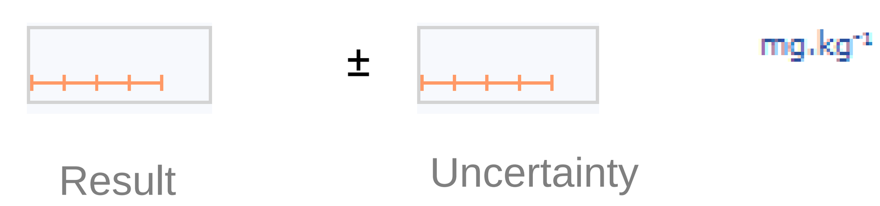
Caroline LAURENT[1], Eric ZIEGLER[1], Abdelkader BOUBETRA[1], Anne TIRARD[1]
[1] : Bureau Interprofessionnel d'Études Analytiques (BIPEA) - 189 rue d'Aubervilliers, 75018 PARIS – France. Tel .+33 1 40 05 26 30
Corresponding author: eziegler@bipea.org

INTRODUCTION

In order to compare the results of different determinations or to compare a result with a threshold or a regulatory value, it can be useful to have information on how the considered results are reliable. The level of confidence in the result can be expressed as the uncertainty around the result's value. Be able to report the measurement's uncertainty together with the result of an analytical determination is a requirement for laboratories accredited according to ISO/IEC 17025 standard [1]. Several ways to calculate uncertainties are possible and procedures to follow are often not that clear and easy for laboratories, statistics being not in their core of business. Some methods can be found in different guides and normative documents (e.g. ISO 21748 [2], Nordtest, GUM). The evaluation of how relevant are these calculated uncertainties is not obvious. Therefore, proficiency tests (PT) can be a useful tool to evaluate how much realistic the estimated uncertainties are, and allow the comparison of the different uncertainties provided by the laboratories.

IMPLEMENTATION OF UNCERTAINIES IN BIPEA'S PROFICIENCY TESTS

Bipea runs almost one hundred regular proficiency testing scheme (PTS), in various areas, such as food, beverages, cereals, soils, water, fragrances, sunscreen products and so on. According to the need of the participants the request of uncertainties can be implemented at each round, once a year or not requested at all. The uncertainty for the concerned parameters can be filled on the usual online reply form in a dedicated associated box, just besides the results, and can be filled in or not. (example on the right).



EXPLOITATION OF THE DATA

First of all, the choice made by BIPEA is to have a pedagogical approach, and therefore to provide information about the uncertainty to its participants but not to give any judgment about it, the proficiency remaining evaluated in regards of the usual standard deviation for proficiency assessment and expressed as z-score.

For each parameter, the minimum and maximum uncertainties limits, u_{min} and u_{max}, are calculated according to ISO 13528:2015 standard [3], which defines u_{min} as equal to the uncertainty associated to the assigned value $u(X_{pt})$, and u_{max} as equal to 1.5 times the robust standard deviation of the results. With an expanded coefficient k=2, U_{min} = 2 x u_{min} and U_{max} = 2 x u_{max}. The laboratory can first consider if its provided uncertainty is within the range [U_{min} ; U_{max}] or not.

A graphical representation is then provided, showing the laboratories performance expressed as z-score, as a function of the provided uncertainty. An example of such graph is provided below (cf. figure1)
It allows to quickly see the distribution of the provided uncertainties (along the abscissa axis), if the provided uncertainties are within [U_{min} ; U_{max}] and the distribution of the results according to the provided uncertainty.

area corresponding to true results AND relative uncertainty "under control".

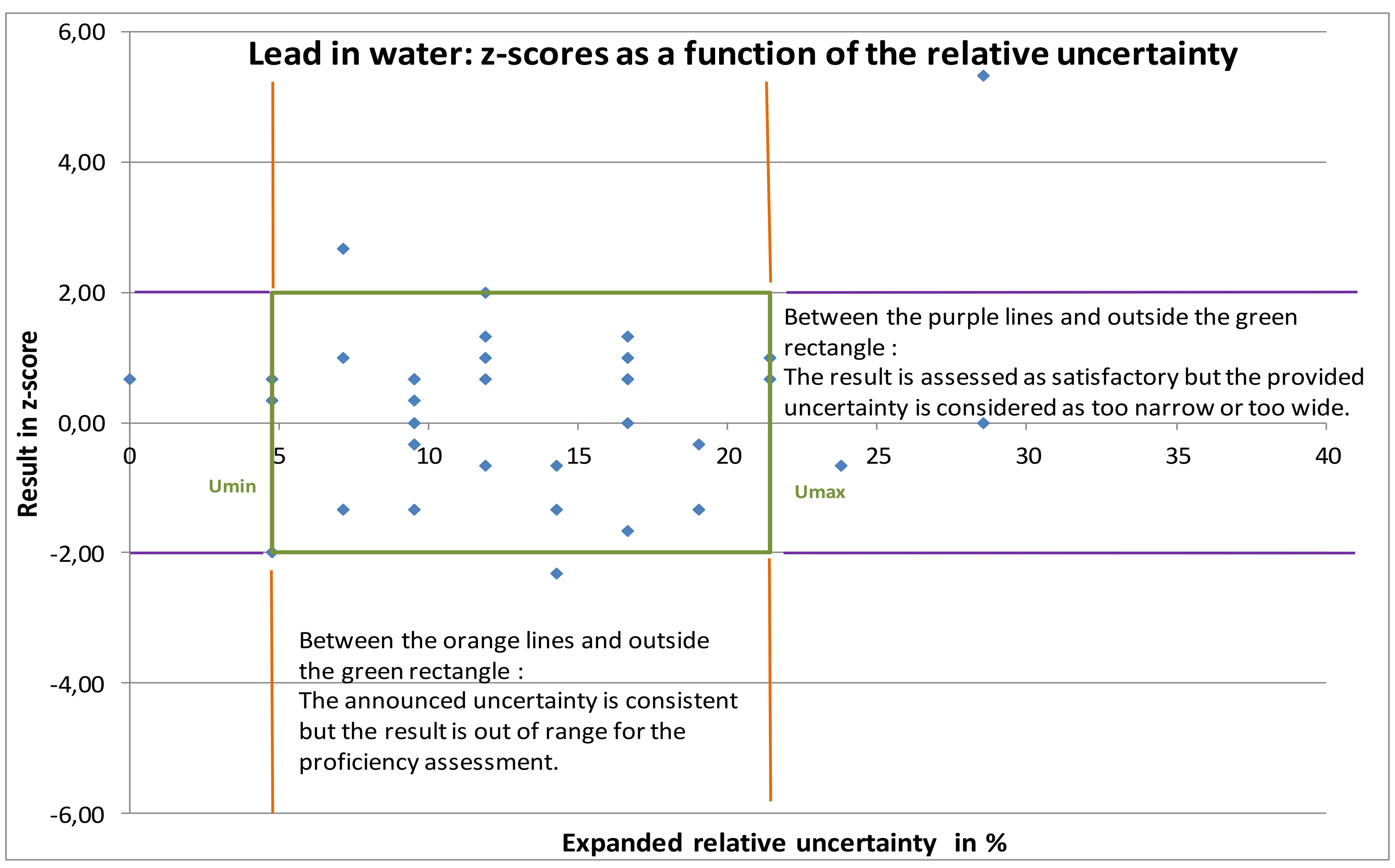


Fig. 1 – Example of representation of the z-score as a function of the uncertainty

ZETA-SCORE

Zeta-score is a usual tool to evaluate if the provided result is an accordance with the assigned value in regards of the announced uncertainty. Zeta-score is calculated as follows:

$$\zeta_i = \frac{x_i - x_{pt}}{\sqrt{u^2(x_i) + u^2(x_{pt})}}$$

Where
X_i is the result of the participant
X_{pt} is the assigned value
$u(X_i)$ is the uncertainty of the determination provided by the laboratory
$u(X_{pt})$ is the uncertainty associated to the assigned value.

However some restrictions have to be considered:
- ISO 13528:2015 standard [3] states that zeta-scores are not so appropriate when the assigned value is the consensus mean of the participants, which is often the case.
- an overestimation of the uncertainty leads to good zeta-scores
- different assessments can be difficult to understand by the participants, and some can for example argue that they have a good zeta-score in spite of a out of range z-score, and consequently considered this result as satisfactory.

That is why BIPEA decides not to carry out an assessment from the zeta-scores. Zeta-score can however be provided but for information purpose only. It can remain a useful tool, especially by following it through time to evaluate if the estimated uncertainty is appropriate or not:
- if the average zeta-score value (mean of the absolute value) is low, the estimated uncertainty may be too large
- if too many zeta-scores are out of range the estimated uncertainty is probably too reduced.

RESULTS and DISCUSSION

First of all, the response rate concerning the uncertainty is very diverse from one PTS to another and ranges from 20% to 60%.
Secondly, the uncertainty provided for a same parameter varies a lot from a laboratory to another (ratio from 1 to 6 in the example in the graph), even when similar methods or techniques are used. Consequently, the main difference comes from the estimation of the uncertainty, that is to say what is taken into account in the estimation (one or several matrices, operators, devices, etc.) and how the calculations are performed. Requesting the uncertainty can, from this point of view, allows the participants to see what is announced by the others and if their uncertainty is in the average or on the contrary quite different from the others.
The graphs like in the example above define big areas in which the laboratory can situate its result. Between the two purple lines, the result is satisfactory but the announced uncertainty is considered as too narrow (on the left of the green rectangle) or too wide (on the right of the green rectangle). Between the two orange lines, the announced uncertainty is in the expected range but the result is too far away from the assigned value. The green rectangle, defining the area where both the result is satisfactory and the uncertainty suitable, is often quite wide, especially regarding U_{min} and U_{max}, but some laboratories remain however outside of it.
It can then often be noted that the performances obtained in the PTS are not linked to the announced uncertainty. As shown in the example of the graph, the distribution of the z-scores is quite flat in regards of the announced uncertainty. Most with a high announced uncertainty have results as good as the others and on the contrary some have out of range results (|z-score| >2) despite announcing low uncertainties.
The use of such graph seems therefore interesting to bring information to the participants, without evaluating them, and so encourage them to take part to this additional section dedicated to the uncertainty.

CONCLUSION

The request to the participants of an interlaboratory test to provide the uncertainties associated to their results can be very useful whatever their state of advancement on this subject. It allows to raise awareness to inexperienced laboratories and allow to evaluate their uncertainties for the more advanced ones. From the collect of the uncertainties, it can first be noted that for a same determination the range of the provided uncertainties is often quite wide, and secondly that the performances observed in the test are very poorly related to the level of the uncertainties announced in the test.

REFERENCES

[1] **ISO 17025** – General requirements for the competence of testing and calibration laboratories.
[2] **ISO 21748** – Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty evaluation.
[3] **ISO 13528** - Statistical methods for use in proficiency testing by interlaboratory comparisons.