

The opinions expressed in the following article are entirely those of the authors and do not necessarily represent the views of either The Royal Society of Chemistry or the Editor of *The Analyst*.

Quality Issues

The
Analyst

The Concept of Uncertainty as Applied to Chemical Measurements

William Horwitz* and Richard Albert

Food and Drug Administration, HFS-500, Washington, DC 20204, USA

The calculation of uncertainty as recommended for physical measurements cannot be transferred readily to chemical measurements. Physical measurements and chemical measurements have entirely different error patterns that behave differently on replication. Correctable local bias predominates in physical systems and random error is minor; random error predominates in chemical systems and bias is difficult to identify and eradicate. Therefore bias must be monitored by randomizing in the interlaboratory environment, a concept not handled by the conventional ISO treatment of uncertainty.

Keywords: *Uncertainty; measurements; error; variability; reproducibility; interlaboratory study*

Measurement uncertainty is a concept that provides an interval within which a true, exact and correct value is expected to lie. This term evolved within the past decade from efforts to avoid overemphasizing the reliability of measurements made to establish fundamental national and international measurement standards. These primary measurement standards are transferable through a logical and physical progression to other laboratories so that measurements taken at different times and at different places are comparable. All of the labor expended by scientists and engineers throughout the world would be useless without the intercomparability and traceability of our measurement system.

The concept of uncertainty was transferred to chemists of the national standards institutions by the assignment of absolute values to certified reference materials (CRMs). CRMs are carefully prepared materials that contain chemical entities in exact and constant concentrations. The certifying laboratory exerts considerable effort to ensure homogeneity and to minimize bias and random error in the measurements. The final objective is to encompass within a range an estimate of where the 'true' value exists with a high degree of confidence, typically 95 or 99%. This halo of uncertainty serves a valuable purpose to provide a narrow reference interval that will be transferred through subsequent chemical operations in other laboratories to estimates of values for those same chemical entities in similar materials.

In this progression of operations, chemists encounter their first exposure to statistical operations; the law of propagation of error, where the effect of small errors pile on top of one another not merely by simple addition but by vector addition. The result is an amplification of that original circle of uncertainty into a larger circle, but one that was usually well within the specifications.

The Transfer of the Concept of Uncertainty from Physics to Chemistry

A problem develops because a rigid structure that is useful in analyzing and transferring physical measurements is imposed upon the completely different error pattern of chemical measurements.

Physical measurements are applied largely to properties involving mass, length, time, temperature, and charge. Many factors that need correction merely require adding or subtracting a constant to adjust for locally constant environmental influences such as pressure, temperature, and gravity. The measured parameters often contain from five to nine significant figures, which change according to physical laws expressed in finely tuned equations. Experimental error in fundamental physical measurements has been whittled down to almost negligible proportions.

Initially, the concept of uncertainty as applied to CRMs in analytical chemistry appeared to be successful. The assigned concentrations were reliably specified in the sense of magnitude, but also allowed for any uncertainty that arose from the intrinsic displacements of physical measurements and from the inherent variability in the chemical operations. Furthermore, these standards were also intended to be the reference points for national standards that 'facilitate intercalibration of laboratories and comparabilities of measurements.'¹ In the area of physical measurements, the values assigned by the primary laboratory were intended to be the reference points required to calibrate the secondary and tertiary generations of laboratories. Every time a new value was generated, however, the uncertainty in the measured value 'was increased due to uncertainties in the laboratories own standards and propagation of the uncertainty of measurements.'¹

A primary purpose of CRMs was to provide an independent external monitor of the effectiveness of the statistical control of laboratory measurements. A fundamental distinction exists between routine laboratory measurements conducted in the day-to-day operations of the laboratory and the long-term capabilities of the chemical measurement system (CMS). Routine measurements provide the laboratory output; the capabilities are reflected in the control chart limits for the laboratory operations. If these limits are too broad for the purpose at hand, the capabilities must be enhanced. The day-to-day values are also applicable to the occasional nonroutine examinations from the laboratory if due allowance is made for the sporadic nature of the analyses.

The uncertainties associated with the long-term CMS parameters consist of the random measurement fluctuations and systematic laboratory biases. These two types of deviations have fundamentally different statistical properties: random errors add vectorially but systematic errors add algebraically.

What this means to the analyst is that systematic errors are all on an equal footing, contributing directly according to their magnitude and sign. With random errors, however, a larger than typical error can have an overwhelming influence on the final error budget. The problem of mixtures of systematic and random errors was discussed by Hansen *et al.*,² particularly in conjunction with small samples, as is frequently the situation encountered in analytical chemistry. These authors compared the difference of the mean square error (the average square of the deviations of the estimates from the true value) from the variance (the square of the deviations from the average of the estimates). The ratio of the two types of errors is a measure of the relative influence of the factors.

The significance of this point lies in the nature of the errors constituting the physical and chemical domains. With physical systems, systematic errors predominate. These are 'corrected out of the result' to arrive at a value recognized as a physical constant or standard, verifiable independently in different laboratories. Biases are not necessarily randomly distributed and are independent of the number of measurements. With chemical systems, normally distributed random errors predominate; typically their magnitude is inversely related to concentration, C . This random error has been derived as a function of the resolution, dC/C , of the CMS³ and can be expressed in terms of a simple exponential expression:

$$\text{RSD}_R (\%) = 2C^{-0.1505}$$

where $\text{RSD}_R (\%)$ is the relative standard deviation among laboratories, C is in mass/mass units ($100\% = 1.0$), and 0.1505 is $0.5 \log_{10} 2$. Some typical values of this random error of the CMS taken from food analysis are given in Table 1. The effect of random error can be decreased by increasing the number of measurements.

The 20–50% standard uncertainties arising from chemical operations in the typical laboratory examining food for contaminants today are definitely not the uncertainties contemplated by the originators of the uncertainty concept. Note also that when the standard uncertainty approaches the 50% point, the system goes out of control because the expanded uncertainty (twice the standard uncertainty) encompasses zero, thereby introducing the overriding additional uncertainty of the absence of the analyte.

A single value for a constant, systematic error is sufficient to provide a correction term. A single value for a random error term, however, is 'floating' in measurement space and can only be anchored to some extent by producing additional concentration estimates under reasonably constant conditions. The amplitude of the random error domain varies inversely with the square root of the number of replications performed. Conse-

quently the uncertainty of such a series of measurements is at the mercy of your financial budget, not your technical budget or professional skill.

The Trapdoor Convenience

In the development of the uncertainty concept, the generators recognized that error budgets could not anticipate all possible deviations that might be uncovered in the future. Moreover an underlying assumption is made that the measurements are in statistical control. They provided themselves with what the lawyers call a trapdoor, a mechanism that ordinarily lies unused, but which allows convenient escape when necessary. When the uncertainty concept was applied to analytical chemistry by EURACHEM,⁴ the trapdoor concept proved to be a convenient way to handle major errors that were not foreseen by these experts. Two examples of these 'Type B' guesstimated uncertainties included the allowance for chemical problems in the acid–base titration example and the estimate based on the experts' opinion of the repeatability of measurements of the injection of pesticide residues into the chromatographic column. In the titrimetry case, the effect of CO_2 on the endpoint overrides all other sources of measurement error; in the residue case, the error budget calculation grossly exaggerated the actual uncertainty estimates. The fact that the parent acid, although labeled similarly, is supplied in different strengths led to overlooking this important source of interlaboratory random error in the leaching of toxic elements from ceramic ware. In the final analysis, the attempts to provide valid estimates of chemical uncertainty by budget analysis proved fruitless.

There are so many potential sources of error in chemical measurements that a requirement to set up an error budget is almost preordained to failure. The quest for the origin of perturbations in final values in chemical operations is a never ending search, leading at best to new PhDs in analytical chemistry and at worst to the expansion of the chemical literature with 'new and improved' methods of analysis. In the absence of CRMs, the progression of improvements can introduce unsuspected biases in the process so that the 'improved' final result is a long way from the 'true value'. Such is the status of carbohydrate analysis, which includes such ill-defined components as polysaccharides, fiber, starch, lignin, *etc.*, in food analysis. Note should be taken of these 'self-defining methods of analysis'. For these there are no external reference standards because the result itself is the 'true' concentration by definition.

The Top-down Approach

Perhaps sufficient time has not elapsed for the approach of the Analytical Methods Committee (AMC)⁵ to permit an overall estimate of the uncertainty from interlaboratory studies to percolate into the uncertainty discussion. The primary advantage of this methodology is that it is based on estimates from actual laboratory measurements, not from expert guesstimates and budget allowances. It is an overall value that uses laboratories as a random factor (a sample of laboratories is used), not confined to the single-laboratory, single-instrument restriction of the uncertainty documents (where the laboratory is a fixed factor). Subtle mistakes such as correcting for the same deviation twice in different contexts are avoided. Best of all, this proposal eliminates the bureaucratic tendency to spend dollars to save pennies. The AMC proposal is a more realistic, all-inclusive approach that has the advantage of over 100 years of experience in evaluating method performance, easily extended to laboratory performance. With a little statistical manipulation, these interlaboratory parameters can be generalized as uncertainty intervals and plotted as the control chart limits that are now generally accepted in the modern laboratory.

Table 1 Expected variability among-laboratories* at various concentrations (order of magnitude) with typical examples from food analysis

Concentration	RSD among-laboratories* (%)	Examples
100%	2.0	Sugar, salt
10%	2.8	Protein in flour
1%	4.0	Salt in food
100 mg per 100 g	5.7	Calcium in milk
10 mg per 100 g	8	Vitamin C in fruits
1 mg per 100 g	11	B vitamins in bread
1 mg per kg (ppm)	16	Pesticide residues
10 µg per kg	32	Aflatoxins, vitamin A
1 µg per kg (ppb)	45	Drugs as tissue residues
≈ 0.5 µg per kg (5E-10)	≈ 50	Practical limit of reliable measurement

* Within-laboratory RSD is expected to be about a half to two thirds of these values..

The major advantage of the top-down approach is that it randomizes the locally constant individual laboratory biases into interpretable critical limits that include the major sources of chemical deviations. These are laboratories, analysts, methods, and time-factors that are left out of the uncertainty error budget calculations. The ISO document on uncertainty calculations did not consider measurements beyond the single instrument stage as practicable.⁶ Observations not obtained by repeated observations but evaluated by scientific judgement are relegated to a 'Type B' status. This fuzzy category is unnecessary in the top-down approach. The top-down proposal supplies values that can be used almost directly for statistical control of chemical measurements within- and among-laboratories.

We suggest that instead of trying to disentangle the various threads involved in the error budget approach to uncertainty, let the measurements speak for themselves. Why bother with the various individual sources of bias and imprecision when what the chemist wants to know is merely the final integrated result. To probe the separate components of error is a noble undertaking that may enable a researcher to optimize a given chemical method of analysis. But that would be a very special case. Once an optimized method is at hand, the user need not generally be concerned with the so-called error budget. Many of the individual factors contributing to variability of chemical measurements are intrinsically unknowable anyhow. Yet sound statistical investigation can provide an effective handle to predict the overall effect of these hidden variables. The situation is analogous to the radiochemist not knowing what particular nuclei of ²³⁵U will decay within a given time, even though it is guaranteed that a definite fraction of those nuclei will decay. We only care about the net result, not the individual components that contribute to it.

Conclusion

Without a refinement of concepts, the metrologists risk losing a large part of their chemical constituency. The presentations of the metrologists suffer from a lack of clarity and transparency to a chemical audience. Their failure to understand chemical variability leads them to inflate some variabilities (specific, local measurements that are later also incorporated into more encompassing measurements) and overlook the more important sources of chemical variability (interferences and effect of concentration differences). A more meaningful approach would replace the metrological ISO within-laboratory uncertainty by the more comprehensive analytical chemistry concept of among-laboratory reproducibility, the randomized individual-laboratory biases combined with the pooled within-laboratory variability.

References

- 1 Taylor, J. K., *Handbook for SRM Users*, NBS Special Publication 260-100, National Institute of Standards and Technology, Gaithersburg MD 20899, 1985.
- 2 Hansen, M. H., Hurwitz, W. N., and Madow, W. G., *Sample Survey Methods and Theory*, Wiley, New York, 1953, p. 58.
- 3 Albert, R. and Horwitz, W., *Anal. Chem.*, 1997, **69**, 789.
- 4 *Quantifying Uncertainty in Analytical Measurements*, Eurachem Secretariat, PO Box 46, Teddington, Middlesex TW11 0LY, UK, 1995.
- 5 Analytical Methods Committee, *Analyst*, 1995, **120**, 2303.
- 6 International Organization for Standardization, *Guide to the Expression of Uncertainty in Measurement*, ISO, Geneva, 1995, p. 53.