



## Proficiency testing of sampling†

Cite this: *Anal. Methods*, 2017, 9, 4110 Analytical Methods Committee AMCTB No. 78

Received 23rd June 2017

DOI: 10.1039/c7ay90092a

rsc.li/methods

Proficiency testing of analytical laboratories is now ubiquitous—laboratories need to participate regularly to receive accreditation. The outcome helps participants to detect unexpected sources of uncertainty. But sampling, as well as analysis *per se*, introduces its own uncertainty in most types of analysis. Sampling uncertainty arises partly from heterogeneity in the ‘target’ (the particular mass of material for which the sample is a surrogate), but also from variation in the manner in which the sample is extracted. Different samplers will do the job somewhat differently, even when following a single protocol. The potential value of proficiency tests for sampling is therefore obvious. And in recent years, feasibility studies and schemes have appeared in several application sectors. Different sectors have very different constraints on how sampling and proficiency testing can be carried out. Nevertheless, the need for general guidelines is obvious.

## Basic limitations in a sampling proficiency test (SPT)

The well-established pattern for operating analytical proficiency tests is a useful starting point for considering what a basic layout for sampling proficiency testing might look like. In analysis the usual pattern is this: participant laboratories are sent a portion of the effectively-homogeneous test material, and asked to determine specific analytes, within prescribed uncertainty limits, by using any procedure, method, or measurement principle. The results are converted into readily-understandable scores. We need to do something similar with sampling, but a moment’s thought tells us that it’s going to be much more complicated. Consider the following.

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7ay90092a

- We can’t send the target to the participants: nearly always the samplers will have to travel to the target.
- The samplers usually can’t have one target each. They will all have to use the same target sequentially, which implies that, to preserve independence, the target must be capable of being restored to its original appearance and, possibly its original constitution, between each sampling episode. For the same reason, the samplers should not see each other at work.
- The target must be acceptably close to stable in composition over the period in which the samplers operate.
- The target will not necessarily be effectively homogeneous. In extreme cases, differences among samplers could be overwhelmed by the heterogeneity of the target.
- Unless all of the samplers use the same detailed protocol, it is impossible to separate between-protocol variance and between-sampler variance (this applies equally to analysis).
- The composition of each sample has to be determined by analysis, which adds further component of uncertainty to the outcome. Moreover, there are two options for the analysis: having all of the analysis undertaken in one laboratory, or each sampler separately commissioning the analysis of their own sample.
- The target will in most instances comprise a valuable commodity, and owners will not want its value to be reduced by the exercise.

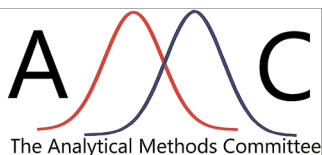
Phew! These difficulties, however, were recognised from early considerations of uncertainty from sampling<sup>1–3</sup> and have been addressed in various ways. In newly emerging applications, or where some doubt exists about the validity of the sampling protocol, some preliminary validation would be required before proficiency testing could be usefully undertaken. That could be tackled by a series of randomised replicated experiments to estimate the variance components associated with different targets, protocols, samplers, samples and analysis. The protocol could then be suitably tailored to match the qualities of the target material and the requirements of fitness for purpose (unless it was a protocol prescribed by a regulator).

While the different sources of variation in results might be untangled by a randomised replicated experiment, such designs

## AMC Technical Briefs

are produced by the Analytical Methods Committee (AMC) - the branch of the Analytical Division that handles matters of technical importance to the analytical science community.

Find out more at: [www.rsc.org/amc](http://www.rsc.org/amc)



do not lend themselves to use in a proficiency test, where a participant needs a score analogous to the z-score,  $z = (x - x_{pt})/\sigma_{pt}$  derived from a result  $x$ , an assigned value  $x_{pt}$ , and a fitness-for-purpose criterion  $\sigma_{pt}$  that is independent of the results.

## Heterogeneity and the sampling protocol

In a practicable sampling proficiency test, the reported result will inevitably reflect the heterogeneity of the target as well as the proficiency of the sampler. This is seldom a problem in applications where an established sampling protocol would be used by all participants. In such instances the protocol would have been specifically designed to accommodate the characteristics of the typical target and thus reduce the effect of heterogeneity to a manageable level.

Even in some established application areas, however, there is no uniformly accepted protocol. In such instances the results will inseparably reflect the variation among the protocols, as well as in the performance of the samplers (including their skill in choosing the most appropriate protocol).

## Sampling and analytical uncertainties

Once the primary samples have been collected, there are alternative ways in which the analysis can be commissioned. A simple option is for all of the samples to be both reduced to test samples and then analysed under repeatability conditions, that is, in one run in a single laboratory. A high-precision analytical procedure would allow the variation among the samples to be most obvious. Indeed, if the reduction/analytical standard deviation were less than about one third of the between-sample standard deviation, it might have no detectable effect at all on the z-score. The fitness criterion could then address the sampling uncertainty alone.

A different approach involves each participant organisation providing its own analytical result. That in effect regards the sampling and analysis as a unitary measurement operation for determining the composition of the target. In that circumstance, the fitness criterion should address the combined uncertainty (that is, sampling plus analytical).

## The assigned value and the fitness criterion

The most practicable route to determining the assigned value in sampling is usually finding a consensus from the participants' results. That involves estimating the location of the results (often a robust mean). A shortcoming of this approach is that, at present at least, there would usually be a small number  $n$  of participants and therefore a noticeable correlation between an individual result and the assigned value. This

correlation has the effect of reducing the dispersion of z-scores by a factor of  $(1 - 1/n)$ : they would tend to be misleadingly small although the discrepancy would be negligible for  $n > 15$ . Participants would seem to score somewhat better than reality. It would therefore be appropriate (and, with varying  $n$ , more consistent round-to-round) to modify the z-score to  $z_n = (x - x_{pt})/(\sigma_{pt}\sqrt{1 - 1/n})$ .

In some media (gases in particular), an alternative assigned value can be obtained by spiking the test material with a known concentration of the analyte. In that case the assigned value can be determined independently of the participants' results, so the question of correlation does not arise.

## Experience to date

The first published realisation of an SPT was in 1995, measuring heavy metals in contaminated land as an example to demonstrate the feasibility of the concept.<sup>3</sup> Since then SPTs have been applied to various analytes in workplace air, soil, landfill gas, wheat, green coffee, lettuce, butter, apple juice, stack gas and waste water. Most of these were one-off feasibility studies for research purposes. Some, however, have been used on a regular basis (details of these studies are given in ESI).

**M. H. Ramsey (University of Sussex) and M. Thompson (Birkbeck University of London).**

*This Technical Brief was prepared by the Subcommittee for Uncertainty from Sampling and approved by the Analytical Methods Committee on 06/06/17.*

## References

- 1 M. H. Ramsey, Error estimation in environmental sampling and analysis, in *Sampling of environmental materials for trace analysis*, ed. B. Markert, VCH, Weinheim, 1994, pp. 93–108.
- 2 M. Thompson and M. H. Ramsey, Quality concepts and practices applied to sampling – an exploratory study, *Analyst*, 1995, **120**, 261–270.
- 3 A. Argyraki, M. H. Ramsey and M. Thompson, Proficiency testing in sampling: pilot study on contaminated land, *Analyst*, 1995, **120**, 2799–2804

**CIPD Certification** I certify that I have studied this document as a contribution to Continuing Professional Development.

Name.....

Signature.....Date.....

Name of supervisor.....

Signature.....Date.....