Measurement Uncertainty –
Comparing the GUM and 'top down' approaches

Upon requests, I tabulate the differences and advantages / disadvantages of the two broad approaches in measurement uncertainty (MU) evaluation processes.

GUM (bottom up) approach	Top down approaches
Component-by-component using Gauss' error propagation law for uncorrelated errors	Component-by-component using Gauss' error propagation law for uncorrelated errors
Which components? Studying uncertainty contributions in each step of test method as much as possible	*Which components?* Using repeatability, reproducibility and trueness of test method, according to basic principle: *accuracy = trueness (estimates of bias) + precision (estimates of random variability)*
"Modeling approach" or "bottom up approach", based on a comprehensive mathematical model of the measurement procedure, evaluating individual uncertainty contribution as dedicated input quantities	"Empirical approach" or "top up approach", based on whole method performance to comprise the effects from as many relevant uncertainty sources as possible using the method bias and precision data. Such approaches are fully in compliant with the GUM, provided that the GUM principles are observed.
Acknowledged as the master document on the subject of measurement uncertainty	There are few alternative top down approaches, receiving greater attention by global testing community today
GUM classifies uncertainty components according to their method of determination into type A and type B: Type A – obtained by statistical analysis Type B – obtained by means other than statistical analysis, such as transforming a given uncertainty (e.g. CRM) or past experience	Top down approaches consider mainly Type A data from own statistical analysis from within-lab method validation and inter-laboratory comparison studies

GUM assumes that systematic errors are either eliminated by technical means or corrected by calculation.	The top down approaches allow for method bias in uncertainty budget
In GUM, when calculating the combined standard uncertainty of the final test result, all uncertainty components are treated equally	The top down approach strategy combines the use of existing data from validation studies with the flexibility of additional model-based evaluation of individual residual effect uncertainty contributions.
Advantages: 1. Demanding critical assessment and full understanding of the analytical steps in a test method 2. Consistent with other fields of measurements such as calibration 3. The MU result generated is relevant to the particular laboratory that produces it	**Advantages**: 1. Quality data from method validation and inter-lab comparison studies are readily available in a well run accredited laboratory 2. Very much simpler process in MU evaluation 3. The MU data of a test method is dynamic and current, due to using existing and experimentally determined quality control checks and method validation results 4. This approach is based on statistical analysis of data generated in intra- and inter-laboratory collaborative studies on the use of a method to analyze a diversity of sample matrices.
Disadvantages: 1. The GUM approach process is tedious and time consuming 2. This methodology may underestimate the measurement uncertainty, partly because it is hard to include all possible uncertainty contributions 3. GUM may unrealistically assume certain errors are	**Disadvantages:** 1. The top down approach may not by itself identify where the major errors could be occurring in process and the results generated are the products of technical competence of the laboratory concerned 2. That inter-lab reproducibility data considered in certain

random (i.e. normally distributed) and independent 4. GUM provides a broad indication of the possible level of uncertainty associated with the method rather than a measurement. 5. It does not take into account either matrix-associated errors or the actual day-to-day variation seen in a laboratory 6. GUM does not apply well when there is no mathematical model in the test method	instances may not be fully representative for variability of results on actual samples, unless it is standardized