Katarzyna Grebieszkow

“New” way of uncertainty calculation
based on
Guide to the expression of uncertainty in measurement (GUM)

According to Bureau International des Poids et Measures (BIPM) already in 1979
“almost all [laboratories] believed that is was important to arrive at an internationally accepted procedure for expressing measurement uncertainty and for combining individual uncertainty components into a single total uncertainty”.

NA61 and NA49 Collaboration Meeting
8-12 X 2012 CERN

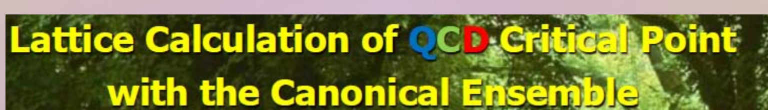
In 1995 International Organization for Standardization (ISO) published „Guide to the expression of uncertainty in measurement” (GUM) - a document describing how to calculate and express uncertainties.

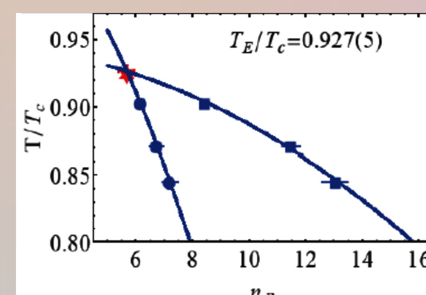
So far we were ignoring it....

Notation $\Phi_{pT} = 7.6 \pm 0.6$ (stat.) ± 1.2 (sys.) MeV/c was used and fully understood in scientific communities.

However recently I have noticed:

$T_{CP} = 0.927(5)\ T_C$ (~ 157 MeV)
$\mu_{CP} = 2.60(8)\ T_C$ (~ 441 MeV)





In the “new” method two kinds of errors (now called uncertainties!) are unified (both expressed as standard uncertainties/deviations) and then added to obtain final standard uncertainty/deviation (see next pages). Notation “±” is now commonly used for expanded uncertainty which is typically 2-3 times standard uncertainty!!

Uncertainty (of measurement) - parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

Standard uncertainty (here I denote it as u_x, in Guide denoted as u(x)) – uncertainty of the result expressed as a standard deviation (for example, the standard deviation of the mean).

Type A evaluation of uncertainty - a method of calculating measurement uncertainty by the statistical analysis of a series of measurements. This method can be based on each correct method of statistical analysis of data. Examples: calculating standard deviation of the mean, using method of least squares to fit a curve to the data and then calculating the parameters of the curve and their standard uncertainties.

Type B evaluation of uncertainty - a method of calculating measurement uncertainty using ways other than the series of measurements, or by a method other than type A. This kind of calculating (or rather estimating!) uncertainty is usually based on scientific judgment of the investigator taking into account all available information, which may include: the results of previous measurements, experience and knowledge of the behavior and properties of instruments and materials, the manufacturer's information on multimeter properties and precision, etc.

Important: **Errors** (exact values unknown and unknowable) ≠ **uncertainties** (can be evaluated)

True value – a value that would be obtained by a perfect measurement (cannot be determined).

Error (of measurement) – result of a measurement minus a true value of the measurand. For example **systematic error** – mean that would result from an infinite number of measurements minus a true value of the measurand (systematic error and its causes cannot be completely known).

We should try to correct our results for systematic errors (at least for recognized effects). The combined (see later) standard uncertainty of corrected result should include both the uncertainty of the uncorrected result and uncertainty of the correction.

GUM and this presentation is focused on uncertainties.

☹ Even if the evaluated uncertainties are small, there is no guarantee that the error in the measurement is also small. A systematic effect may have been overlooked because it is unrecognized. Finally, our corrections may be imperfect.

☺ The (corrected) result of the measurement can unknowably be very close to the true value of measurand (and hence have a negligible error) even though it may have a large uncertainty.

The uncertainty of the result of the measurement should not be confused with remaining unknown error (see back-up slides for graphical illustration of true value, errors, and uncertainties).

$$u_x(\text{type A}) = s_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N \cdot (N-1)}}$$

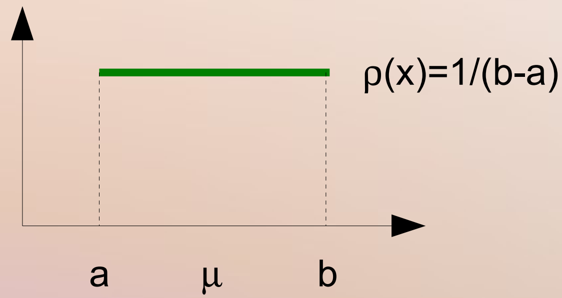
← <u>example</u> of uncertainty calculation type A (experimental standard deviation of the mean); applied for series of measurements when probability distribution of x_i is given by Gaussian distribution

$$u_x(\text{type B}) = \frac{\Delta x}{\sqrt{3}} = \sqrt{\frac{(\Delta x)^2}{3}}$$

or

$$u_x(\text{type B}) = \sqrt{\frac{(\Delta x)^2}{3} + \frac{(\Delta x_E)^2}{3}}$$

← the most popular <u>example</u> of type B; here probability distribution of x_i is given by uniform (rectangular) distribution



Expected value $\mu = (a+b) / 2$;
Variance: $Var = (b-a)^2 / 12$
$a = -\Delta x$ and $b = \Delta x$ => $Var = (\Delta x)^2/3$
finally: standard deviation $u_x = \sqrt{Var}$

Δx – called sometimes uncertainty limit / boundary uncertainty / calibration uncertainty. Due to finite resolution of used devices/meters. But sometimes we have additionally:

Δx_E – uncertainty of experimenter in reading analog meters display ("parallax error"). We can neglect it if we believe that we can read out everything very precisely.

$$u_x(\text{total}) = \sqrt{u_x^2(\text{type A}) + u_x^2(\text{type B})} = \sqrt{s_{\bar{x}}^2 + \frac{(\Delta x)^2}{3} + \frac{(\Delta x_E)^2}{3}}$$

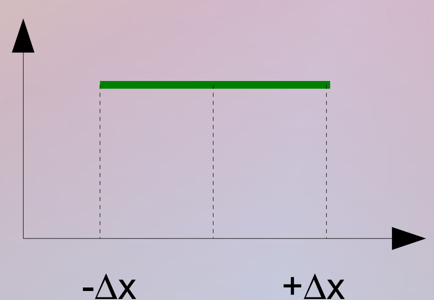
total uncertainty of a direct measurement

Note: both types of uncertainty evaluation **(type A and B) are** similar and **based on probability distributions. The uncertainty components** resulting from either type **are quantified by** variances or **standard deviations.**

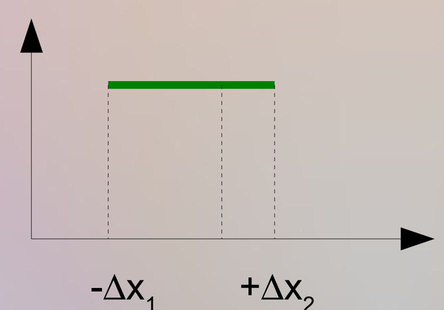
Type A standard uncertainty is obtained from a **probability density function** derived from an **observed frequency distribution** (the statistical distribution of the results of series of measurements).

Type B standard uncertainty is obtained from an **assumed probability density function** based on experience of other information (often called subjective probability).

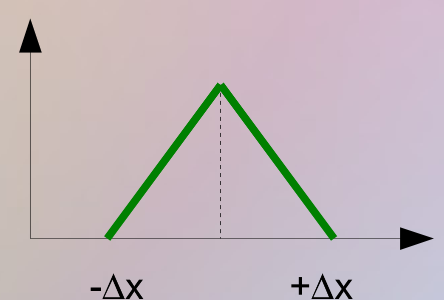
Examples (see GUM for trapezoidal, etc. distributions):



Most popular
rectangular distribution
standard deviation $u_x = \Delta x/\sqrt{3}$



asymmetric distributions are also discussed – see GUM F.2.4.4 and G.5.3



triangular distribution
standard deviation $u_x = \Delta x/\sqrt{6}$

Combined standard uncertainty (uncertainty propagation)

We have a quantity y, which is a combination of **independent(!)** a, b, and c, where a, b, and c are measured with total(!) uncertainties u_a, u_b, and u_c. Then the combined uncertainty u_y (in GUM denoted as $u_c(y)$) can be taken from the *law of propagation of uncertainty*.

$$y = f(a, b, c)$$

$$u_y = \sqrt{\left(\frac{\partial f}{\partial a}\right)^2 u_a^2 + \left(\frac{\partial f}{\partial b}\right)^2 u_b^2 + \left(\frac{\partial f}{\partial c}\right)^2 u_c^2}$$

When input quantities a, b, and c are correlated the above formula has to be modified, and, in principle, we have to calculate covariance matrix → see GUM 5.2 for details.

In most cases (including our scientific research) it is enough to report measured value x with its (combined) standard uncertainty and to write it typically as **x (u_x) [unit]**. However, for particular applications it may be need to give also:

Expanded uncertainty (here denoted as U_x, in Guide U) - a value determining the size of the interval around the result of the measurement, which covers a large fraction of the distribution of values that could reasonably be attributed to the measurand. This uncertainty is used to compare the results with results of other experiments, with tables of constants, etc. It is also used for commercial purposes and to establish industry, health, and safety standards.

$\mathbf{U_x = k\, u_x}$ k - coverage factor

The choice of factor k, which is usually in the range **2 < k < 3**, is based on *coverage probability* or *level of confidence* (denoted as p) required of the interval $x - U_x$ to $x + U_x$
In most cases (including our students' laboratories) k = 2 is recommended.

GUM recommendation: **the coverage factor k is always to be stated**, so that the standard uncertainty u_x can be recovered for use in calculating the combined standard uncertainty of other quantity.

The value k=2 determine the probability of finding the real (true) value within the range $x \pm U_x$ as equal to 95% (in case of uncertainty type A) or 100% (for uncertainty type B).
In fact, in case of uncertainty type B p = 100% is reached already for $k = \sqrt{3} = 1.73$.

Example 1 (type B evaluation of uncertainty):
We measure (only once) the length of the table $L = 50$ cm, resolution: ΔL (max) = 1 mm
Note(!): in this example I (intentionally) neglect uncertainty of an experimenter but one might use additionally the estimated value ΔL_E (minimal !) = 0.5 mm

Old method: $L = 50.0 \pm 0.1$ cm (here the estimated uncertainty was a systematic one)

New method: standard uncertainty/deviation type B: $u_L = \Delta L/\sqrt{3} = 0.058$ cm
example of expanded uncertainty: $U_L = 0.1$ cm (for k=1.73)
(for a single measurement with uncertainty type B and PDF rectangular there is no sense to use $k > \sqrt{3}$)

Guide suggests several methods of writing final result:

Final notations

In case of (combined) standard uncertainty:
$L = 50.000$ cm, $u_L = 0.058$ cm or $L = 50.000$ cm, $u_L = 0.58$ mm
L = 50.000 (58) cm ← widely used in scientific publications, catalogs, tables, etc.
L = 50.000 (0.058) cm
$L = (50.000 \pm 0.058)$ cm ← formally allowed but NOT recommended by GUM (our students are not allowed to use it!) because if may be confused with expanded uncertainty.
GUM reminds that ± should be used to indicate an interval corresponding to a high level of confidence.

In case of <u>expanded</u> uncertainty:

L = (50.0 ± 0.1) cm (k=1.73), but if possible we should also write **p (here 100%),** the way how k was selected (for example PDF normal / **PDF rectangular**), number of degrees of freedom **ν** (see Annex G for details), etc.

other examples:

L = (50.000 ± 0.096) cm (k=1.65), p=95%, PDF rectangular

L = (50.000 ± 0.099) cm (k=1.71), p=99%, PDF rectangular

p - *coverage probability* or *level of confidence* of the selected interval. Whenever practicable, the level of confidence associated with the interval defined by U_x should be estimated and stated. However it can be difficult because it requires the detailed knowledge of the probability distribution (complicated if we have e.g. both type A and type B uncertainties contributing to combined uncertainty; convolution function of probability should be obtained first).

The whole Annex G of GUM discusses how to select the value of k, that produces an interval having a level of confidence close to a specified value. However, **for many practical measurements in a broad range of fields**, when standard uncertainties $u_{x,i}$ (which may be obtained from either type A or type B evaluations) contribute comparable amounts to the **<u>combined</u> standard uncertainty** we can (with a help from Central Limit Theorem):

1. **adopt k=2 and assume that p ≈ 95%**
2. **adopt k=3 and assume that p ≈ 99%**

More about writing final results...

Two rules which are, of course, not changed:

1. Both **the value and its uncertainty should be written with the same precision.**

$L = 50.000$ cm, $u_L = 0.058$ cm

even if expressed in different units:
$L = 50.000$ cm, $u_L = 0.58$ mm

2. The **uncertainties can have maximum 2 significant digits** (although in some cases it may be necessary to retain additional digits to avoid round-off errors in subsequent calculations).

Example 2 (types A and B evaluation of uncertainty):
50 measurements of a diameter (x) of the pencil by use of the micrometer screw (resolution of the screw gives maximally $\Delta x = 0.01$ mm, estimated uncertainty of experimenter while reading is minimum $\Delta x_E = 0.005$ mm)

New method:

Measurements [mm]: 6.25, 6.25, 6.27, 6.22, 6.23, 6.23, ... (50 measurements)

$$x \equiv \bar{x} = \frac{1}{50}\sum_{i=1}^{50} x_i = 6.26\,[\text{mm}]$$

standard uncertainties/deviations:

$$u_x(\text{type A}) \equiv s_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^{50}(x_i - \bar{x})^2}{50\cdot(50-1)}} =$$

$$= \sqrt{\frac{1}{2450}[(6.25-6.26)^2 + (6.25-6.26)^2 + (6.27-6.26)^2 + \ldots]} = 0.0028 \quad [\text{mm}]$$

$$u_x(\text{total}) = \sqrt{u_x^2(\text{type A}) + \left(\frac{\Delta x}{\sqrt{3}}\right)^2 + \left(\frac{\Delta x_E}{\sqrt{3}}\right)^2} =$$

$$= \sqrt{0.0028^2 + 0.0058^2 + 0.0029^2} = 0.0071 \quad [\text{mm}]$$

Examples of final notation:
x = 6.260 (0.007) mm or
x = 6.260 (7) mm or
x = (6.260 ± 0.014) mm (k=2)

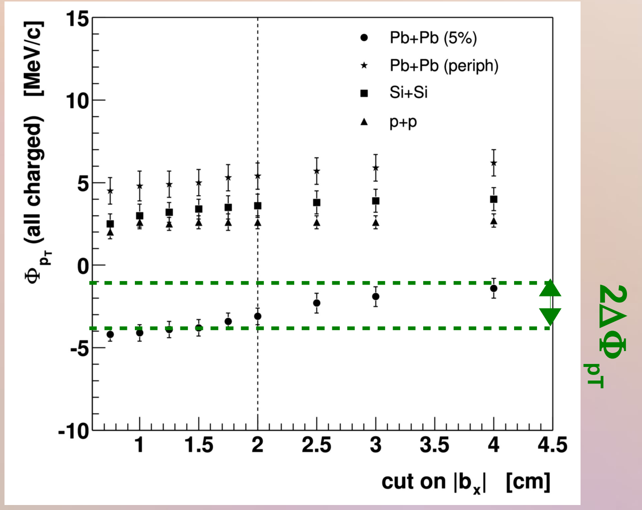
Example 3 (types A and B evaluation of uncertainty):

Old method: Φ_{pT} = 7.6 ± 0.6 (stat.) ± 1.2 (sys.) MeV/c

New method: total (A and B types) standard uncertainty:

$$u_{\Phi_{pT}} = \sqrt{0.6^2 + \frac{1.2^2}{3}} = \sqrt{0.36 + 0.48} = 0.916 \ \ \text{MeV/c}$$

Note: here uniform (rectangular) distribution assumed in case of 1.2(sys.) uncertainty (1.2 treated as Δx_E).
But GUM describes also triangular, trapezoidal, etc. PDF *).



thus finally:

Φ_{pT} = 7.60 (0.92) MeV/c or

Φ_{pT} = 7.6 (0.9) MeV/c or

Φ_{pT} = 7.60 (92) MeV/c or

Φ_{pT} = (7.6 ± 1.8) MeV/c (k=2), etc.

*) type B analysis of uncertainty is sometimes cumbersome, but it leads to better understanding of the experiment and sometimes helps to optimize it

Example 4 (combined uncertainty):
Determination of a resistance R
We measured (only once):

U = 26 V using analog multimeter (range 0-30V; multimeter class: 1)
ΔU = class x range /100 = 1/100 x 30 V = 0.3 V

I=0.825 A using digital multimeter (range 2A; for this range ΔI =1.2% x rdg + 1 dgt)
ΔI = 1.2% x 0.825 A + 1 x 0.001 A = 0.0109 A

New method: standard uncertainties/deviations type B:
$u_U = \Delta U/\sqrt{3} =$ 0.173 V => U = 26.00 (0.17) V (here we neglect "parallax error")
$u_I = \Delta I/\sqrt{3} =$ 0.00629 A => I = 0.825 (0.006) A

$$\text{combined uncertainty } u_R = \sqrt{\left(\frac{\partial R}{\partial U}\right)^2 u_U^2 + \left(\frac{\partial R}{\partial I}\right)^2 u_I^2} = \sqrt{\left(\frac{1}{I}\right)^2 u_U^2 + \left(\frac{-U}{I^2}\right)^2 u_I^2} = 0.319\,\Omega$$

R = 31.51 (0.32) Ω or
R = 31.51 (32) Ω or
R = (31.51 ± 0.64) Ω (k=2), etc.

Old method: R = U/I ± ΔR, where

$$\Delta R = \left|\frac{\partial R}{\partial U}\right| \Delta U + \left|\frac{\partial R}{\partial I}\right| \Delta I \qquad \text{(total differential method)}$$

or

$$\frac{\Delta R}{R} = \frac{\Delta U}{U} + \frac{\Delta I}{I} \qquad \text{(logarithmic differential method)}$$

CANNOT BE USED ANY MORE !!!

Because there is no separate method for "systematic error" (or rather uncertainty) estimation. For each measured value (here U and I) we should first calculate their (total) standard uncertainties (u_U, u_I) and then propagate these total standard uncertainties to obtain combined standard uncertainty of R.

Summary: I do not intend to convince you to use the new method... But it is worth to know and understand this new notation.

“New” method of uncertainty calculation:

☺ is unified (common way of treating both types of uncertainties A and B); evaluation of measurement uncertainty is fully consistent and transferable

☺☹ results in only one final value of uncertainty

☹ can be sometimes problematic (in a few years from now younger students may think that 0.6 value in $\Phi_{pT} = 7.6 \pm 0.6$ MeV/c means for example 2 times standard uncertainty, it is expanded uncertainty, instead of standard uncertainty)

Guide: *„3.4.8 Although this Guide provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value.”*

1. I am very grateful to A. Kubiaczyk (WUT) for his useful comments

2. GUM can be taken from:
http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

3. Some (fragments of) definitions were rewritten directly from GUM

Backup slides

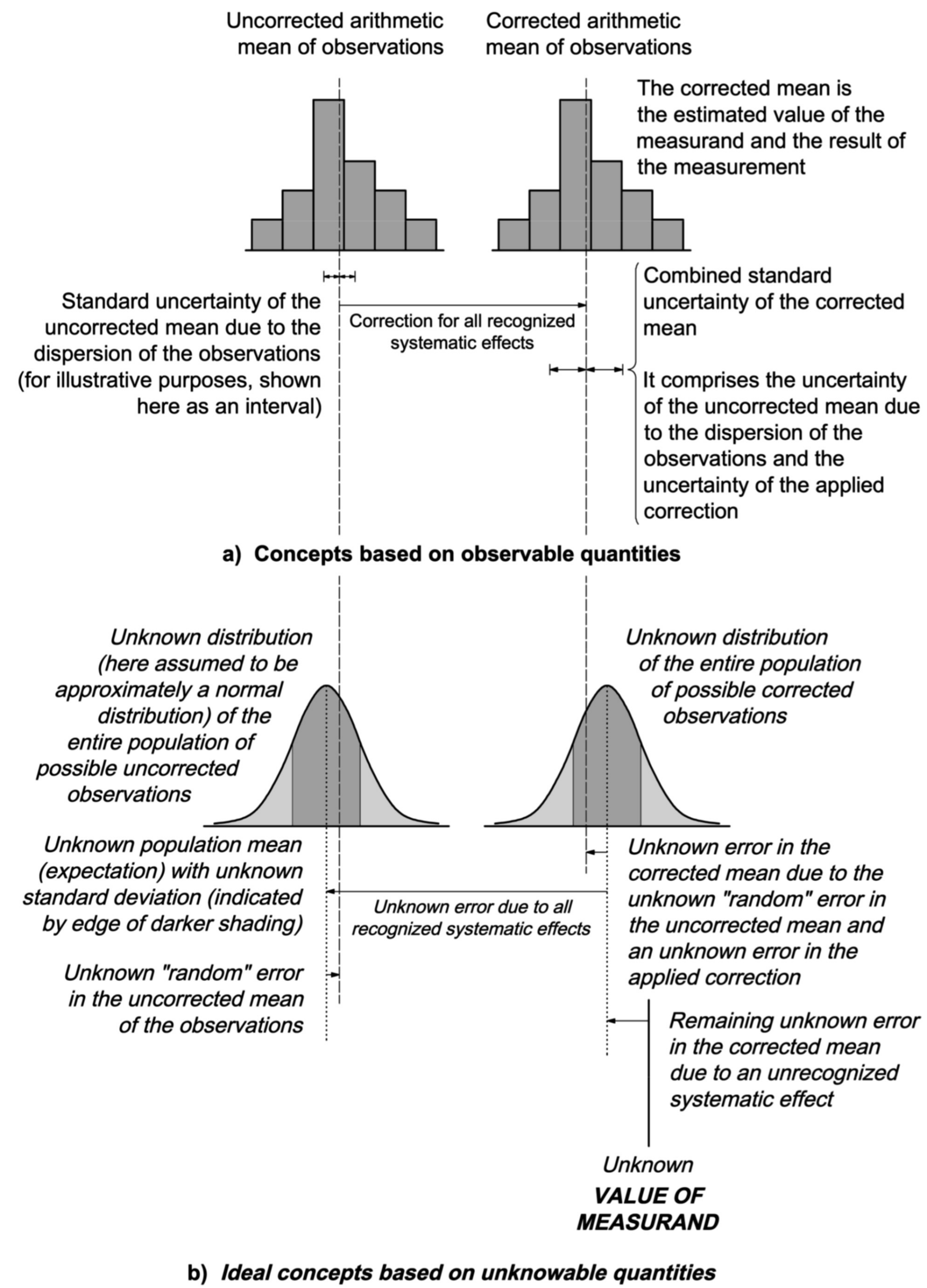


Figure D.1 — Graphical illustration of value, error, and uncertainty

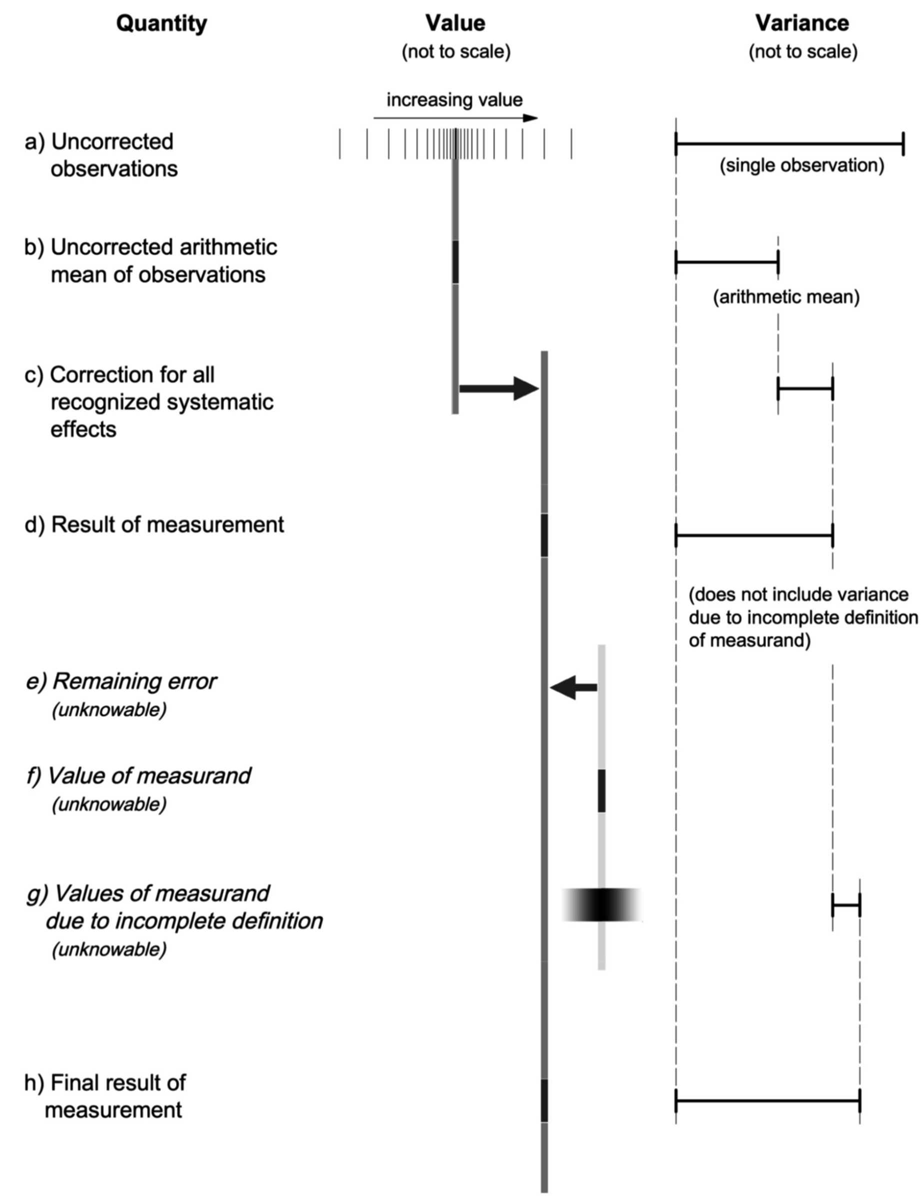


Figure D.2 — Graphical illustration of values, error, and uncertainty