JOURNAL OF CHEMICAL EDUCATION

pubs.acs.org/jchemeduc

Article

Integrating Measurement Uncertainty Analysis into Laboratory Education for the Development of Critical Thinking and Practical Skills

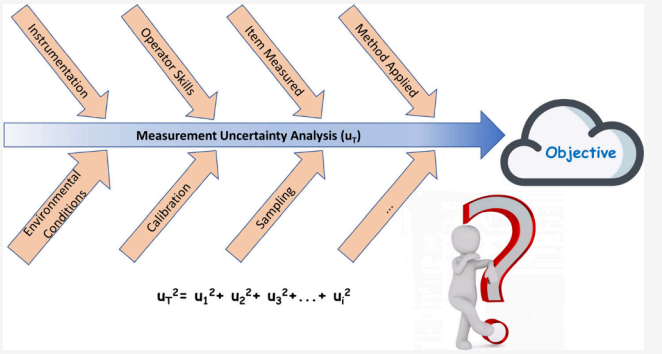
Juan M. Sanchez*

Cite This: J. Chem. Educ. 2024, 101, 4783−4789

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Laboratories that have successfully passed an accredited laboratory program need to validate their methods, a requirement that should extend to all types of testing laboratories. Such validation includes the performance of uncertainty analysis. However, the teaching and practice of concepts of measurement uncertainty analysis remain underemphasized in most science degrees. In this study, a laboratory session is proposed to introduce students to the in-house calibration of a laboratory balance. It provides valuable hands-on experience for students to understand measurement uncertainty analysis, particularly in the context of balance calibration. By involving students in calculating uncertainty and critically evaluating the relevance of their results, the session not only teaches specific skills but also nurtures important aspects of scientific thinking. A bottom-up approach is employed in which students use their knowledge to assess the significance of information. This proves to be an effective method for promoting critical thinking, encouraging students to actively engage with the material and apply their judgment, which is crucial for developing their scientific reasoning abilities. Furthermore, by incorporating concepts such as the minimum weight for accurate measurements, practical relevance is added to the exercise, helping students understand the real-world implications of uncertainty analysis in experimental work. By bridging theoretical concepts with practical applications, students gain a deeper appreciation of the importance of uncertainty in scientific research.



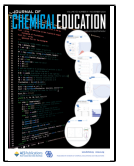
KEYWORDS: *Upper-Division Undergraduate, Interdisciplinary/Multidisciplinary, Decision Making, Chemometrics*

Accuracy has been defined as the "closeness of agreement between a measured quantity value and a true quantity value of a measurand."[1] In this common definition, the concept is independent of the precision. However, the "true value" of a measurand in experimental sciences is never known, and it is not possible to give a quantity value for accuracy. For this reason, more recent usage considers accuracy to be a combination of two measures: trueness (or bias) and the precision of a method, and it understands that an accurate method requires both good precision and good trueness.[2−4] Following this new usage, precision must be considered when determining accuracy.

The initial step in developing an analysis method should be to identify the accuracy goals (trueness and precision) of the proposed method and determine the equipment required to meet these standards. All testing methods should be validated before application to ensure that trueness and precision meet established criteria and to ascertain the uncertainty associated with the method's application. Only in this way can accurate measurements be ensured. While bias can be eliminated if no systematic errors take place or if they are corrected, measurements always carry some degree of uncertainty that can never be entirely eliminated. The International Vocabulary of Basic and General Terms in Metrology (VIM) defines uncertainty as a "non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand."[1] In line with this, all laboratories should be mindful of the uncertainty in their measurements given that it is inherent in scientific methodologies, concepts, and communication of science.[5,6]

In undergraduate science education, it is important to explain to students that uncertainty is an essential concept in laboratory measurements. Measurements are never conducted under perfect conditions, and even the most advanced laboratories have their limitations.[7] When identical conditions and equipment are used for repeated tests, the results may not be consistent.[2] Hence, "no measurement can ever be perfectly precise, and uncertainties help us understand the range within

Received: May 15, 2024
Revised: September 27, 2024
Accepted: September 30, 2024
Published: October 10, 2024



https://doi.org/10.1021/acs.jchemed.4c00583
J. Chem. Educ. 2024, 101, 4783−4789

which the value of a measurement is likely to fall."[8] Furthermore, measurement uncertainty (MU) analysis is a requirement for all laboratories that want to pass an accredited laboratory program and should be incorporated into all testing procedures. Students need to grasp that accurately reporting an experimental result with its uncertainty is necessary to determine whether the result is fit for purpose.

For many years, calculating the standard deviation associated with a series of replicate measurements using a method has been considered a simple procedure for determining the uncertainty associated with that method.[9] However, the uncertainty of a measurement method is influenced by numerous experimental factors, including the measuring instrument and materials used, the item being measured, environmental conditions, operator skills, calibration standards, and the sampling procedure.[10] Thorough understanding and knowledge of the entire measurement method are therefore needed in order to be able to conduct a correct uncertainty analysis.

As uncertainty is dependent on the method and instrumentation employed, uncertainty analysis must be performed for each specific method and defined sample type during method validation. This analysis provides an estimate of the greatest variability that may reasonably be expected for that particular method under the conditions used, and the calculated uncertainty may be applied to all determinations that are so described. When uncertainty analysis is conducted correctly, variabilities larger than the stated uncertainty should rarely occur when applying the validated method. However, it is essential to remember that if any parameters of the measurement method or the sample type are altered, the uncertainty analysis must be reassessed and adjusted accordingly.

MU analysis is not a simple task, and consequently, the subject is widely misunderstood; people may even feel daunted by it.[6,7] It is likely that this is the main reason why MU analysis is seldom explained to undergraduate chemistry students and why it appears in so few analytical or physical chemistry textbooks.[11,12] When the topic is introduced, it is typically explained solely from a theoretical perspective or in the context of performing error propagation calculations based on given data.[13−23]

This study introduces a teaching methodology designed to help students understand the importance of MU in laboratory analysis. The methodology initially provided students with a brief introduction to the statistical equations necessary for basic MU calculations. Students practice how to apply critical reasoning to assess whether the equipment available in the laboratory to conduct their experiments may significantly impact the final measured uncertainty. Additionally, the methodology shows how to consider the stated precision requirements when selecting the most suitable instrumentation.

BRIEF INTRODUCTION ABOUT APPROACHES TO MEASUREMENT UNCERTAINTY ANALYSIS

There are several different approaches to conducting uncertainty analysis, each tailored to the specific requirements and characteristics of the measurement process. Regardless of the approach used, the goal of uncertainty analysis is to provide a quantitative assessment of the reliability of measurement results, enabling scientists to reach informed decisions and conclusions based on their data.

One approach is the Monte Carlo method,[24,25] which involves simulating the measurement process using random sampling techniques. In this approach, the measurement process is mathematically modeled using computers, and random variations in input parameters are generated to simulate real-world conditions. By repeating this process numerous times, a distribution of possible measurement outcomes is obtained, allowing for the estimation of uncertainty. This is a powerful and robust technique, which has traditionally been difficult to integrate into laboratory sessions due to a lack of availability of simple tools to carry out the simulations.[26] However, the availability of web-based tools today, such as the NIST Uncertainty Machine, makes the Monte Carlo method far more widely available.[17]

There are two other approaches for evaluating measurement uncertainty in laboratory analysis: the bottom-up approach (or the Guide to the Expression of Uncertainty in Measurement, GUM, approach),[27] which is a modeling or metrological approach, and the top-down approach (experimental approach), which uses validation and control quality data (such as the Nordtest methodology).[28]

The top-down approach is particularly practical for testing laboratories, as it is relatively easy to use and cost-effective.[29] However, it requires uncertainty to be estimated from the evaluation of a large amount of experimental data, derived from preliminary quality control and method validation data,[30,31] making it time-consuming. Unfortunately, the limited time available for experimental sessions in laboratory lessons makes it difficult to properly apply this approach.

The bottom-up approach (GUM) involves a systematic evaluation and estimation of the individual contribution of each step of the method to the overall uncertainty.[27] It requires a thorough understanding and study of the method to estimate all potential sources that could impact the uncertainty, which can be challenging in laboratory analyses. The approach's reliance on complex statistical procedures is considered to be a limitation.[32,33] Additionally, it has been criticized for being less applicable for evaluating uncertainty in the results of testing laboratories.[30] However, the GUM approach remains a useful approach for metrological laboratories, instrument calibrations, and reference material certification.[10] Moreover, when testing laboratories lack sufficient experimental data to establish a reliable estimate of uncertainty using the top-down approach, or when method performance data is unavailable, the bottom-up approach is the most suitable procedure.[10] One of the significant advantages of the GUM approach is its high teaching value, as it fosters critical thinking by requiring the use of available a priori knowledge and leaving the evaluation of the relevance of available information to the scientific judgment of the operator.

In the GUM approach, two types of uncertainties are considered.[10,27] One arises from the variability in repeated determinations of the measurand (type A uncertainty), which assumes a normal probability distribution for the obtained data. However, there are other factors beyond repeated measurement that may contribute to uncertainty, requiring the determination of a second type of uncertainty (type B). Type B uncertainty is an estimation that is specific to the measurement process and requires prior knowledge to identify all potential sources of uncertainty that may have an impact on the method's result. It is important to note that type B uncertainties are nonstatistical contributors to uncertainty that need relatively complex calculations based on probability distribution functions to determine the standard uncertainty value associated with each source. Once calculated, standard uncertainties from each source are treated as individual inputs that contribute to the overall uncertainty in the measurement. Therefore, all individual uncertainties must be combined to derive an overall figure.[7]

https://doi.org/10.1021/acs.jchemed.4c00583
J. Chem. Educ. 2024, 101, 4783−4789

Broadly speaking, type B uncertainty can be understood as the "inherited" variability of the method due to various sources.[7,32]

The analysis of type B uncertainties in the bottom-up approach holds significant educational value, as it requires a preliminary understanding of the method to identify its critical stages. Consequently, the analysis of type B uncertainty can be incorporated into laboratory lessons to teach students how to apply critical reasoning in selecting the most suitable equipment for laboratory work.

PRELIMINARY STATISTICAL CONSIDERATIONS REQUIRED FOR UNCERTAINTY ANALYSIS

Students typically lack formal training in MU analysis. In most cases, there is a need to introduce them to or reinforce the main concepts of MU analysis to ensure that they grasp its significance.

In experimental measurements, it is essential to have values for the uncertainty in each independent source (variable) that can significantly affect the final uncertainty. When a variable is obtained from repeated measurements (repeatability), its uncertainty (type A) is determined from the standard deviation calculated for repeated measurements of the same item. For example, the repeatability uncertainty of a balance can be obtained from repeated measurements of a reference weight. However, if we only have a single value for a variable, its uncertainty (type B) is typically estimated from the information provided by the manufacturer (e.g., tolerance of volumetric material and the uncertainty of a reference material) and the readability of the measurement instrument.

Type B uncertainty analysis requires a priori knowledge of the probability model associated with each variable,[27] and this may vary for each variable. Once the probability functions are known, the variability of each source (e.g., tolerance value given by the manufacturer) must be transformed into an equivalent standard uncertainty using specific calculations.[27,34] The most common probability functions applied for laboratory analyses are normal, triangular, and rectangular (uniform) distributions.[11,27,34] The default distribution used when the actual distribution is unknown is rectangular distribution, which assigns equal probability to finding a value anywhere between the given limits $+a$ and $-a$ and zero probability outside this range.

The standard uncertainty for a uniform distribution is calculated by[11,27,34]

$$u = \frac{a}{\sqrt{3}} \tag{1}$$

where a can be either the tolerance given by the manufacturer or the readability of an instrument.

After determination of all individual standard uncertainties, the next step is to calculate a combined uncertainty. For the purpose of this study, it is assumed that there are no correlations between the uncertainties of the different components, which is the situation that is usually applicable in assay laboratories.[32] Therefore, since standard deviations (uncertainties) are not additive, variances (squared uncertainties) must be used to calculate the combined standard uncertainty:

$$u_T^2 = \sum u_i^2 \tag{2}$$

or

$$u_T = \sqrt{\sum u_i^2} \tag{3}$$

where u_i are the uncertainties for each individual source.

The final reported uncertainty (U, expanded uncertainty) must be determined while considering the level of confidence associated with the reported value. The combined u_T in eq 3 is equivalent to a standard deviation[32] and is assumed to follow a normal distribution. Therefore, the combined standard uncertainty (u_T) represents a margin that can be understood as plus or minus one standard deviation, covering approximately 68% of the probability. To increase the confidence level, a coverage factor (k) must be applied to expand the standard uncertainty. The most common factor is $k = 2$, which approximately covers a 95% confidence interval:

$$U = ku_T \tag{4}$$

CASE STUDY: IN-HOUSE DETERMINATION OF THE UNCERTAINTY OF A LABORATORY BALANCE

To begin, students were introduced to the process of performing in-house calibration of a four-digit analytical balance located in the teaching laboratory (Kern & Sohn GmbH, Balingen, Germany, model ACJ 200-4M, with a reproducibility of ≤0.0002 g). This procedure is relatively simple and can be completed within a single laboratory session. The choice of this case study is significant because balances are among the most commonly used instruments in laboratories and the initial step in nearly all assays conducted in a laboratory involves weighing a sample or a standard. It is crucial for students to understand that the uncertainty associated with a balance can significantly impact the accuracy of the final measurement result of a method.

In our teaching laboratory, we have two precision analytical balances: one with a readability of three digits (0.001 g, Kern & Sohn, model IoT-Line Precision Balance 572-30) and the other with a readability of four digits (0.0001 g, model ACJ 200-4M). Before evaluating the uncertainty of the four-digit balance, students were asked to consider which balance they would choose for their measurements. All students answered that either balance could be used as long as the weight to be measured exceeded their respective readabilities (a, limited resolution on the display). When asked which balance would be used to "accurately" weigh 10 mg (0.01 g), all students considered that either balance could be used without having any effect on the final measured weight. It was found that none of the students participating were aware of the importance of balance uncertainty and its impact on subsequent measurements. Therefore, students were also introduced to a methodology for assessing the effect of balance uncertainty on subsequent measurements.

In the case of balances, various sources of uncertainty are usually considered, including repeatability, uncertainty of the reference weight used in the balance calibration, readability, eccentricity, and linearity. However, in the proposed method, eccentricity and linearity were not assessed because these uncertainty sources are specific to balances, and students at this level may not be familiar with them.

Determination of Type A Uncertainty

To determine type A uncertainty, students used a set of Class F1 reference weights (W1100-200, Accuris Instruments) with a certificate that confirmed the weights meet the OIML regulations and provided an expanded uncertainty for each reference weight.

To aid in selecting the reference weight for calibrating the four-digit balance, students were provided with information from pharmacopeias, which recommend using a reference

https://doi.org/10.1021/acs.jchemed.4c00583
J. Chem. Educ. 2024, 101, 4783−4789

weight between 0.1 g and 5% of the balance capacity.[35,36] Based on this guidance, students chose to perform the repeatability test using a 1 g reference weight. The statistics obtained for the repeatability test (n = 18) were as follows (more significant figures than those required were maintained during calculations): mean = 1.000089 g; SD = 0.000123 g.

Given the assumption of a normal distribution for type A uncertainty, the standard uncertainty associated with the mean has to be determined by calculating the standard error of the mean (SEM):

$$u_{\text{repeat}} = \text{SEM} = \frac{\text{SD}}{\sqrt{n}} = 0.000029 \text{ g} \quad (5)$$

Determination of Type B Uncertainties

Students were informed that one source of uncertainty for a balance always stems from the traceability of the reference weight used, which must be obtained from the certified uncertainty provided by the manufacturer of the reference material. Another source of uncertainty arises from the readability of the instrument used to measure the weight. These are general sources that may significantly affect MU analysis, and students clearly understood why these two sources must be considered. Other sources that are specific to electronic balances, such as eccentricity and linearity, were not considered since students had no previous knowledge of these instrumental sources.

Reference Weight Uncertainty (u_{RW}). When students read the information provided in the certificate of the manufacturer of the Class F1 reference weights, they observed that the uncertainties varied for each weight. Consequently, it became evident to them that the standard uncertainty associated with this source changes depending on the reference weight measured.

The expanded uncertainty provided by the manufacturer for the 1 g reference weight measured was U = 0.0001 g, with a coverage factor of k = 2. Since this uncertainty was obtained from a certified expanded uncertainty determined by applying a coverage factor of k = 2, the standard uncertainty must be calculated as follows:

$$u_{\text{RW}} = \frac{U}{2} = 0.00005 \text{ g} \quad (6)$$

Readability Uncertainty (u_{read}). Readability refers to the limited resolution on a display or scale, representing the instrument's ability to display incremental changes in its output value (quantitation noise). For balances, readability variability is typically considered as half of the last significant digit ($a/2$),[10] and the assumed distribution for balances is rectangular.[10,37]

It is important to note that a weighing operation involves two readings (tare and net sample weight), each of which is associated with rounding errors that are the same ($\sigma_{\text{weight}} = \sigma_{\text{tare}} = \sigma$). Simple error propagation gives additive variations for addition and subtraction, which implies that the variance of the weighed amount (σ_{read}^2) must be determined as

$$\sigma_{\text{read}}^2 = \sigma_{\text{weight}}^2 + \sigma_{\text{tare}}^2 = \sigma^2 + \sigma^2 = 2\sigma^2 \quad (7)$$

Therefore, the standard uncertainty due to readability (rectangular distribution) was calculated as follows:

$$u_{\text{read}} = \sqrt{2\sigma^2} = \sqrt{2\left(\frac{\frac{a}{2}}{\sqrt{3}}\right)^2} = \sqrt{\frac{1}{2}\left(\frac{a}{\sqrt{3}}\right)^2} \quad (8)$$

For the four-digit balance (a = 0.0001 g), u_{read} = 0.000041 g.

Total (u_T) and Expanded (U) Uncertainties. Applying eq 3, the total standard uncertainty of the laboratory balance was determined as follows:

$$u_{\text{T}} = \sqrt{u_{\text{repeat}}^2 + u_{\text{RW}}^2 + u_{\text{read}}^2} \quad (9)$$

$$u_{\text{T}} = \sqrt{(0.000029)^2 + (0.00005)^2 + (0.000041)^2} = 0.000071 \text{ g} \quad (10)$$

The expanded uncertainty (k = 2) was calculated as

$$U = 2 \times 0.000071 = 0.00014 \text{ g} \quad (11)$$

The expanded uncertainty obtained was within the limits reported by the manufacturer of the balance (reproducibility ≤ 0.0002 g), which means that the balance passed the in-house verification for precision.

CRITICAL REASONING TO UNDERSTAND THE EFFECT OF UNCERTAINTY: INTRODUCTION TO THE CONCEPT OF MINIMUM WEIGHT

From the previous results, students calculated that the variance due to repeatability ($u_{\text{repeat}}^2 = 8.4 \times 10^{-10}$) corresponds to 24% of the total variance (u_T^2=5.02×10^{-9}), whereas the variance due to readability ($u_{\text{read}}^2 = 1.68 \times 10^{-9}$) accounts for 34% and the variance due to the reference weight ($u_{\text{RW}}^2 = 2.50 \times 10^{-9}$) is 42%. These percentages helped students understand that uncertainty determination with laboratory measurements should not solely rely on repeatability measurements. Other sources can contribute significantly to the uncertainty. In this specific case, the two other sources assessed have a higher contribution to the total uncertainty (42% and 34%) than the repeatability (24%).

Despite the importance of knowing the uncertainty associated with a laboratory balance, students must apply critical reasoning for the selection of a balance in the laboratory, which should not be solely based on the absolute uncertainty calculated. Students need to learn that the selection of a balance must take into account not only the uncertainty of the balance but also the objective and use of the weighed substance, which requires the maximum acceptable percentage of uncertainty for that weight in the applied method to be considered.

For instance, in the preparation of a buffer solution, the final concentration of salts in the buffer typically does not require high precision, and uncertainties of 5−10% in concentration values can be acceptable. This means that an ordinary top-pan balance (with a = 0.01 g and u_{read} = 0.0041 g) can be used to weigh corresponding salts with measured weights as low as 0.10 g, resulting in a precision due to repeatability of 4.1%. However, for weighing a sample to be tested and for preparing a standard solution used for calibration, the precision of the weighed amount must be sufficient to ensure that weighing errors are small or negligible relative to other errors generated in subsequent steps of the analytical method.[35] This makes it necessary to introduce the concept of a minimum weight to be applied for quantitative methods.

Despite weighing being a critical operation in all laboratories, there are few references available about the minimum weight to be measured with a laboratory balance outside of analytical books.[38] Only the European Pharmacopeia[35,39] and the US Pharmacopeia[36,40] have specific monographs on weighing and the significance of the balance uncertainty. Both pharmacopeias consider a precision acceptance criterion of ≤0.1% for a balance

https://doi.org/10.1021/acs.jchemed.4c00583
J. Chem. Educ. 2024, 101, 4783−4789

Table 1. Statistical Results Obtained when Assessing the Trueness and Precision of a Four-Digit Balance

Reference Weight	Mean (n = 10)	SD (n = 10)	RSD	SEM[a]	p-value[b]
0.05 g	0.05004 g	0.00015 g	0.30%	0.000047	0.424
1 g	1.00007 g	0.00013 g	0.013%	0.000041	0.143
200 g	200.00022 g	0.00050 g	0.00025%	0.00016	0.197

[a]Standard error of the mean (SEM), calculated as SD/$\sqrt{n}$. [b]One-sample Student's t test.

to be appropriate.[35,36,38−40] For this reason, the concept of the minimum weight that should be measured for a specific balance is introduced in the pharmacopeias. The equation proposed by the pharmacopeias to calculate the experimental minimum weight is defined as[35,36,38−40]

$$\text{maximum percentage of precision } (X\%) = 100 \times \frac{2u_{\text{weight}}\ (\text{g})}{\text{smallest weight (g)}} \quad (12)$$

$$\text{smallest weight (g)} = \frac{200u_{\text{weight}}\ (\text{g})}{X\%} \quad (13)$$

The constant value 2 is used to apply a coverage factor k = 2 to the standard uncertainty (needed for 95% confidence).

At this point, it was explained to students that the goal in all quantitative determinations is to determine the content of an analyte in a sample with the lowest possible uncertainty. With instrumental determinations, methods used always require a final instrumental measurement of the signal obtained for a treated sample and to interpolate the result obtained in a calibration curve determined with calibration standards that are prepared from a standard solution, which has to be prepared by weighing a pure substance. The most common regression method applied for transforming the measured signal into a concentration value is linear regression based on ordinary least squares. This mathematical model assumes that the independent variable (represented on the abscissa axis) has a variability that is significantly less than that of the dependent variable (i.e., instrumental signal). Therefore, it is necessary to know the common level of precision obtained with the instrument used to determine the maximum tolerance acceptable for the variables on the abscissa axis (concentrations). For chromatographic methods, the accepted precision for repeatability of system sustainability tests is usually set at RSD ≤ 1%[41] or RSD ≤ 2%.[42] However, a more stringent acceptance limit, especially given the precision of modern HPLC autosamplers, might be <0.75% RSD.[42,43]

It is important to remind students that the relative uncertainty associated with calibrators and sample concentrations will always be higher than the uncertainty associated with the weight of a standard or sample due to the additive effect of the different sources during the preparation of the solutions to be measured. Given this, weighed amounts of these substances have to ensure a small relative uncertainty, which means that the precision defined by the pharmacopeias (precision of ≤0.1%) is small enough to reach the most common requirements of quantitative methods.

After reviewing the analyses to be performed in their upcoming experiments, all of which were quantitative determinations, the students concluded that the precision criteria given by the pharmacopeia (≤0.1%) could be applied for weighing their standards and samples. In this situation, eq 13 becomes

$$\text{smallest weight (g)} = \frac{200u_{\text{weight}}}{0.1} = 2000u_{\text{weight}} \quad (14)$$

Therefore, for the four-digit balance used in our laboratory, students calculated that a minimum weight of 0.1420 g was required for weighing their samples and substances for preparing standard solutions.

WHY IS IT IMPORTANT TO ANALYZE AND MINIMIZE UNCERTAINTY?

It is important to examine how changes in input parameters impact a model's output, which helps to identify the most influential factors and quantify their influence. This approach aids in prioritizing resources and efforts by focusing on key variables. Furthermore, reducing uncertainty in measurements has been critical in numerous scientific discoveries. Enhanced precision in measurements fosters greater certainty in our findings.

It is important for students to incorporate these concepts when evaluating testing methods. They need to learn how to use previous knowledge to identify the sources of uncertainty that could substantially influence the final uncertainty of a method. Once these sources are determined, implementing best practices becomes essential in reducing uncertainties in measurements. One such practice is to select equipment with the lowest uncertainties.

After conducting the MU analysis to determine the uncertainty of the four-digit balance, students were asked to identify potential sources of uncertainty for balances that could result in either improved or worsened precision. They quickly recognized that the readability uncertainty depends solely on the type of balance chosen and that this value varies significantly depending on the balance used for weighing. Therefore, students were tasked with calculating the readability associated with different types of balances and explaining the effect that it would have on the weights measured.

Using eq 8, students determined that the readability uncertainty for a three-digit electronic balance (a = 0.001 g) is u_{read} = 0.00041 g (0.41 mg), whereas for a top-pan balance (a = 0.01 g), it is u_{read} = 0.0041 g (4.1 mg). This shows that the readability uncertainty of a three-digit balance is 10 times worse than that of a four-digit balance (u_{read} = 0.000041 g). Consequently, the total uncertainty of the three-digit balance must be at least 10 times higher, significantly impacting the precision reached for each balance.

ASSESSMENT OF ACCURACY (TRUENESS AND PRECISION) FOR THE FOUR-DIGIT BALANCE

To effectively demonstrate to students the significance of uncertainty in accuracy, they were asked to perform an assessment of bias and precision on the four-digit balance. Three reference weights were provided to evaluate the balance's trueness and precision across different weighing levels. One reference weight was selected that was near the maximum capacity of the balance (200 g); another matched the reference

https://doi.org/10.1021/acs.jchemed.4c00583
J. Chem. Educ. 2024, 101, 4783−4789

weight used for the MU analysis (1 g); and a third weight was below the calculated minimum weight (0.050 g).

Trueness was assessed using a one-sample Student's *t* test for each of the reference weights. The results, as shown in Table 1, yielded unbiased outcomes across all three weighing levels (for the purpose of this study, significance was set at 0.05, which means that the results yielding $p > 0.05$ are considered as equivalent). In the case of precision, a target RSD of ≤0.1% was selected. For the 50 mg weight (below the determined minimum weight for this balance), the precision achieved was not acceptable, yielding a precision of 0.30%. Therefore, the balance can only be considered accurate (both unbiased and precise) when it meets the minimum weight requirements.

CONCLUSIONS

The proposed methodology provides students with an opportunity to learn how to calculate uncertainty associated with a simple method, such as weighing with a balance, while considering various sources of variability that are inherent in the method. Through this approach, students can explore the concept of minimum weight and understand its significance in quantitative methods. Students discovered that the minimum weight concept used for quantitative analysis is based on a stated uncertainty for the balance.

Students also recognized that by estimating the individual components contributing to the overall uncertainty of a test result, they can assess whether the equipment used is capable of providing accurate measurements. Additionally, this process helped to identify areas within the test method that could be improved to enhance the accuracy. Overall, the described methodology emerges as a valuable addition to science education, providing students with essential skills and insights into the complexity of measurement uncertainty analysis.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.4c00583.

Activity description, students' handout, and questions for students (PDF)

Excel spreadsheet for calculations (XLSX)

AUTHOR INFORMATION

Corresponding Author

Juan M. Sanchez − *Chemistry Department, University of Girona, 17003 Girona, Spain;* orcid.org/0000-0002-4139-7273; Email: juanma.sanchez@udg.edu

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.4c00583

Notes

The author declares no competing financial interest.

REFERENCES

(1) *International vocabulary of metrology - Basic and general concepts and associated terms (VIM)*, 3rd ed.; JCGM, 2012. https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf/f0e1ad45-d337-bbeb-53a6-15fe649d0ff1 (accessed 2024-05).

(2) *ISO 5725-1:2003: Accuracy (trueness and precision) of measurement methods and results - Part 1: general principles and definitions*; ISO, Geneve, Switzerland, 2003.

(3) Menditto, A.; Patriarca, M.; Magnusson, B. Understanding the meaning of accuracy, trueness and precision. *Accred Qual Assur.* **2007**, *12*, 45−47.

(4) Prenesti, E.; Gosmaro, F. Trueness, precision and accuracy: a critical review of the concepts as well as proposals for revision. *Accred Qual Assur.* **2015**, *20*, 33−40.

(5) Gougis, R. D.; Stomberg, J. F.; O'Hare, A. T.; O'Reilly, C. M.; Bader, N. E.; Meixner, T.; Carey, C. C. Post-secondary science students' explanations of randomness and variation and implications for science learning. *Int. J. Sci. Math Educ.* **2017**, *15* (6), 1039−1056.

(6) Rosenberg, J. M.; Kubsch, M.; Wagenmakers, E. J.; Dogucu, M. Making sense of uncertainty in the science classroom: A Bayesian approach. *Science & Education.* **2022**, *31*, 1239−1262.

(7) Bell, S. In *A beginner's guide to uncertainty of measurement*; Measurement Good Practice Guide No. 11 (Issue 2); National Physical Laboratory, Teddington, UK, 1999.

(8) The significance of Measurement Uncertainties. Process Instrument Solutions. https://www.linkedin.com/pulse/significance-measurement-uncertainties-process-instrument-solutions/ (accessed 2024-04).

(9) Fuller, W. A. In *Measurement error models*; Wiley-Interscience, New York, NY, USA, 2009.

(10) EURACHEM/CITAC Measurement Uncertainty Working Group. In *Quantifying uncertainty in Analytical Measurement*, 3rd ed.; Ellison, S. L. R., Williams, A., Eds.; Eurachem, 2012. https://eurachem.org/index.php/publications/guides/quam (accessed 2024-05).

(11) Harris, D. C. In *Quantitative Chemical Analysis*, 9th ed.; W.H. Freeman and Company, New York, NY, USA, 2016.

(12) Garlan, C. W.; Nibler, J. W.; Shoemaker, D. P. In *Experiments in Physical Chemistry*, 8th ed.; McGraw-Hill Higher Education, New York, NY, USA, 2009.

(13) Gordon, R.; Pickering, M.; Bisson, D. Uncertainty analysis by the "worst case" method. *J. Chem. Educ.* **1984**, *61* (9), 780−781.

(14) Guare, J. Error, precision, and uncertainty. *J. Chem. Educ.* **1991**, *68* (8), 649−652.

(15) Yates, P. C. A simple method for illustrating uncertainty analysis. *J. Chem. Educ.* **2001**, *78* (6), 770−771.

(16) Wampfler, B.; Rösslein, M.; Felber, H. The new measurement concept explained by using an introductory example. *J. Chem. Educ.* **2006**, *83* (9), 1382−1388.

(17) Albert, D. R. Monte Carlo uncertainty propagation with the NIST uncertainty machine. *J. Chem. Educ.* **2020**, *97* (5), 1491−1494.

(18) Pernot, P. Comment on "calculating the confidence prediction limits of a rate constant at a given temperature from an Arrhenius equation using Excel. *J. Chem. Educ.* **2017**, *94* (10), 1399−1401.

(19) Olsen, R. J.; Sattar, S. Measuring the gas constant R: propagation of uncertainty and statistics. *J. Chem. Educ.* **2013**, *90* (6), 790−792.

(20) Clark, T. M.; Cervenec, J.; Mamais, J. The price is right" for your classroom. *J. Chem. Educ.* **2011**, *88* (4), 428−431.

(21) Brooks, S. D.; Gonzales, M.; Farias, R. Using surface tension measurements to understand how pollution can influence cloud formation, fog, and precipitation. *J. Chem. Educ.* **2009**, *86* (7), 838−841.

(22) Gao, R.; Lloyd, J. Precision and accuracy: knowledge transformation through conceptual learning and inquiry-based practices in introductory and advanced chemistry laboratories. *J. Chem. Educ.* **2020**, *97* (2), 368−373.

(23) Cerrato-Alvarez, M.; Frutos-Puerto, S.; Pinilla-Gil, E. Use of uncertainty software as a didactic tool to improve the knowledge of Chemistry students in Analytical method validation. *J. Chem. Educ.* **2024**, *101* (1), 104−112.

(24) *JCGM 101:2008: Supplement 1 to the "Guide of the expression of uncertainty in measurement" - Propagation of distributions using a Monte Carlo method*; JCGM, 2008. https://www.bipm.org/en/committees/jc/jcgm/publications (accessed 2024-05).

(25) Jamroz, B. F.; Williams, D. F. Consistency in Monte Carlo uncertainty analyses. *Metrologia.* **2020**, *57* (6), 065008.

(26) Gardenier, G. H.; Gui, F.; Demas, J. M. Error propagation made easy- or at least easier. *J. Chem. Educ.* **2011**, *88* (7), 916−920.

(27) *JCGM 100:2008: Evaluation of measurement data - Guide to the expression of uncertainty in measurement*; JCGM, 2008. https://www.bipm.org/en/committees/jc/jcgm/publications (accessed 2024-05).
(28) Magnusson, B.; Näykki, T.; Hovind, H.; Krysell, M.; Sahlin, E. In *Handbook for calculation of measurement uncertainty in environmental laboratories (NT TR 537)*, 4th ed.; 2017. https://www.nordtest.info/wp/ (accessed 2024-04).
(29) Dabalus Islam, M.; Schweikert Turcu, M.; Cannavan, A. Comparison of methods for the estimation of measurement uncertainty for an analytical method for sulphonamides. *Food Addit Contam Part A Chem. Anal Control Expo Risk Assess.* **2008**, *25* (12), 1439−1450.
(30) Magnusson, B.; Ossowicki, H.; Rienitz, O.; Theodorsson, E. Routine internal- and external-quality control data in clinical laboratories for estimating measurement and diagnostic uncertainty using GUM principles. *Scand J. Clin Lab Invest.* **2012**, *72* (3), 212−220.
(31) Lee, J. H.; Choi, J.-H.; Youn, J. S.; Cha, Y. J.; Song, W.; Park, A. J. Comparison between bottom-up and top-down approaches in the estimation of measurement uncertainty. *Clin Chem. Lab Med.* **2015**, *53* (7), 1025−1032.
(32) Farrance, I.; Badrick, T.; Sikaris, K. A. Uncertainty in measurement and total error - are they so incompatible? *Clin Chem. Lab Med.* **2016**, *54* (8), 1309−1311.
(33) Farrance, I.; Frenkel, R. Uncertainty of measurement: a review of the rules for calculating uncertainty components through functional relationships. *Clin Biochem Rev.* **2012**, *33* (2), 49−75.
(34) Coskun, A.; Oosterhuis, W. P. Statistical distributions commonly used in measurement uncertainty in laboratory medicine. *Biochem Med.* **2020**, *30* (1), 5−17.
(35) Frisch, K.; Saal, C.; Spieldenner, B. Weighing according to the European Pharmacopoeia: general considerations and new general chapter 2.1.7. Balances for analytical purposes. *Pharmeur Bio Sci. Notes* **2022**, *2022*, 1−9.
(36) US Pharmacopeia. Weighing on an analytical balance, general chapter <1251>. In *USP-NF*; United States Pharmacopeial Convention, Rockville, MD, USA, 2021.
(37) Stears, M. In *INL/CON-10-18643: Balance calibration - a method for assigning a direct-reading uncertainty to an electronic balance*; Idaho National Laboratory, 2010. https://inldigitallibrary.inl.gov/sites/sti/sti/4633182.pdf (accessed 2024-04).
(38) McDowall, R. D.; Burgess, C. A question of balance? Part 2: putting principles into practice. *LCGC Eur.* **2006**, *19* (3), 152−159.
(39) European Pharmacopeia. Balances for analytical purposes, general chapter 2.1.7. In *Ph Eur Suppl. 10.6*; The European Pharmacopeia Commission, Strasbourg, France, 2021.
(40) US Pharmacopeia. Balances, general chapter <41>. In *USP-NF*; United States Pharmacopeial Convention, Rockville, MD, USA, 2021.
(41) US FDA. In *Reviewer guidance, validation of chromatographic methods*; Center for Drug Evaluation and Research, Rockville, MD, USA, 1994. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/reviewer-guidance-validation-chromatographic-methods (accessed 2024-04).
(42) US Pharmacopeia. Chromatography, general chapter <621>. In *USP-NF*; United States Pharmacopeial Convention, 2022. https://www.usp.org/harmonization-standards/pdg/excipients/chromatography (accessed 2024-04).
(43) Dong, M. W.; Huynh-Ba, K.; Wong, A. W. Validation of stability-indicating HPLC methods for pharmaceuticals: overview, methodologies, and case studies. *LCGC North Am.* **2020**, *38* (11), 606−618.

https://doi.org/10.1021/acs.jchemed.4c00583
J. Chem. Educ. 2024, 101, 4783−4789