001 supplement

Analysis of Experimental Reliability   SUPL-001

Prepared by R. F. Schneider (Revised 01/12)

concepts

reliability • precision • accuracy • reproducibility • independent measurement • intrinsic precision • average (= arithmetic mean = mean) • average deviation • absolute value • percent • percent deviation • random errors • systematic errors • homogeneous • heterogeneous • significant figures

An essential part of any quantitative laboratory measurement is an understanding of the **reliability** of the measurement. Powerful statistical methods exist for the analysis of reliability. The exercises conducted in the introductory chemistry laboratory will utilize procedures and apparatus with well-known limits of accuracy and precision. Normally, measurements will be replicated several times in order to establish the reliability of those quantities reported as conclusions of exercises (and on which a substantial part of the earned grade will depend). The number of repetitions of a measurement will typically be relatively small (3–4 repetitions). Simple statistical concepts suffice to represent the reliability of measurements that are replicated small numbers of times.

160

Measures of Reliability

We use two quantities to characterize the reliability of a series of experimental measurements—**accuracy** and **precision**. Accuracy measures the extent to which measurements agree with a known or "true" value. Precision measures the consistency of measurements with each other, independent of their accuracy.

Figure 1-1 demonstrates accuracy and precision. It shows the result of a series of attempts to hit the bull's-eye of a circular target with a bow and arrow. The objective of this exercise (i.e., its known value) is to have each arrow strike the center of the target. Concentric circles on the target mark areas which are successively further from the objective. The distance from the center of the bull's-eye represents the error in an attempt. Archer 1 shows great consistency in the part of the target that the arrows strike, but they are all distant from the center of the target. The data are precise, but not accurate. Archer 2 shows inconsistency in the location of the arrows on the target, but there are about as many of them on the right as on the left and on the top of the target as on its bottom. This archer demonstrates some measure of accuracy, but with low precision. Archer 3 demonstrates both accuracy and precision.

How does accuracy play a role in experiments whose results are *qualitative*? If on the basis of a series of experimental results, a sample is asserted to be potassium iodide, the assertion is accurate (i.e., **true**) if the material was, in fact, potassium iodide. If the material was another substance, the assertion is inaccurate, which in this case means **false**. If the data from which the identity was concluded does not permit distinguishing potassium iodide from sodium iodide, asserting that the substance may be one **OR** the other would be an accurate (i.e., **true**) statement if the substance was actually one or the other.
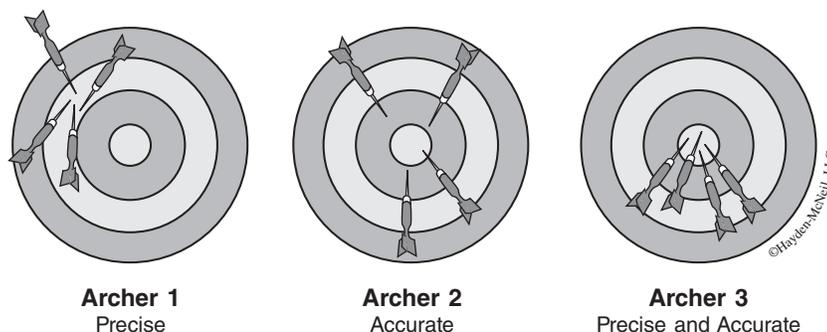

Archer 1
Precise

Archer 2
Accurate

Archer 3
Precise and Accurate

Figure 1-1.

In *quantitative* laboratory exercises, as with the archers, the precision of a result tells us something about how **reproducibly** the procedure was conducted. Are there limits on the precision attainable in laboratory exercises? Yes. Each device and each procedure that is used in an exercise has an **intrinsic** (built-in) **precision**. E.g., the top loading balance can be read to the nearest 0.01 g, but the balance manufacturer guarantees only that if we weigh the same object several times **independently**, (i.e., weigh the object, remove it from the balance, replace it on the balance pan and weigh it again) the indicated weights will not differ by more than 0.03 g. No matter how experienced the individual using the balance may be, the measured weight is uncertain because the balance mechanism introduces small uncertainties. The intrinsic precision of the devices used in an exercise place limits on the overall precision of the exercise.

Quantitative Measures of Reliability

Suppose an exercise calls for the determination of the density of a sample of matter by measuring its volume and mass. The results of two measurements of the same volume result in 31.70 mL and 32.50 mL. What is the "best" value to use for the volume? If the experimenter has no reason to believe that one of the measurements is "better" than the other, the best value to report is the **average** (sometimes called **arithmetic mean,** or simply, **mean**) of the two results, V_{avg} = (31.70 + 32.50)/2 = 32.10 mL. The average of a series of values is defined as the sum of the values divided by the number of the values. I.e.,

$$V_{avg} = \frac{(V_1 + V_2 + \ldots + V_n)}{n}, \text{ or in the case of two measurements, } V_{avg} = (V_1 + V_2)/2$$

Note that the average value has the same **units** as the individual values (i.e., if each volume is measured in mL, the mean will also be in mL).

How shall we measure the **precision** of the two volume measurements? For small numbers of repetitions, a reasonable measure of precision is the **average deviation**. To calculate the average deviation, calculate the difference between each individual measurement and the average (the difference is called a **deviation**), take the **absolute value** of the deviations* (i.e., disregard the sign of the deviation) and average the deviations. For the volume measurements given earlier, the calculation yields:

Avg Dev = [|(31.70 – 32.10)| + |(32.50 – 32.10)|]/2 = [|–0.40| + |0.40|]/2 = 0.40 mL

The final result of the volume determination should be written as

$$V = 32.10 \pm 0.40 \text{ mL}$$

The numerical value of the average deviation reveals a lot about the precision of the volume measurement. However, it is also useful to consider how large the deviation

---

*If we do not take the absolute value of the deviations, the average deviation will always be zero (0). The mean is defined as the number about which the sum of the deviations is zero.

is compared to the quantity being measured. This is accomplished by calculating the **percent deviation** (**percent** means *parts per hundred*) of the measurement, namely,

$$\% \text{ Deviation} = 100 \times 0.40 \text{ mL}/32.10 \text{ mL} = 1.3 \%$$

Note that the units of the measurement appear in both the numerator and the denominator, so that the resulting percent deviation has no physical units.

The second part of a density determination consists of measuring the mass of the object. In this case, three measurements of the same object as above are made resulting in measurements of 41.61 g, 41.72 g and 41.65. The average mass is then:

$$\text{Mass}_{avg} = (41.61 + 41.72 + 41.65)/3 = 41.66 \text{ g}$$

and the average deviation is:

$$\text{Average Deviation} = (0.05 + 0.06 + 0.01)/3 = 0.04 \text{ g.}$$

The mass should be reported as **41.66 ± 0.04 g**. The percent deviation of the mass is:

$$\text{Percent Deviation} = 100 \times 0.04 \text{ g}/41.66 \text{ g} = 0.1\%.$$

What is a quantitative measure of **accuracy**? The experimental result above suggests that the volume is between 32.10 − 0.40 = **31.70 mL** and 32.10 + 0.40 = **32.50 mL**. Suppose the *true volume of the sample* is 31.80 mL. The deviation of the average experimental value from the true value is 32.10 − 31.80 = 0.30 mL. The percent deviation of the experimental result from the true value is $100 \times 0.30/31.80 = 0.94 \%$.

When exercises result in numerical values, grades will generally depend on the **percent deviation of the reported result** as a measure of the **precision** with which the exercise has been conducted. We will often use the **percent deviation of the mean of reported results from the true value** as a measure of **accuracy**.

Computing with Uncertainties

The density computed from the average mass and volume above is d = 41.66 g/32.10 mL = 1.299 g/mL. What precision should be reported for the computed density? While computing with uncertainties can be treated rigorously, we use a simple approach in this illustration. We examine the extremes that result from using the worst combinations of the mass and volume. The *largest* computed density will be that using the *largest mass* and the *smallest volume*, i.e.,

$$d_{max} = 41.70/31.70 = 1.315 \text{ g/mL}$$

The *smallest* density will result from using the *smallest mass* and the *largest volume*, i.e.,

$$d_{min} = 41.62/32.50 = 1.281 \text{ g/mL}$$

The mean of these two values is 1.298. The average deviation of these two values of the density is 0.017 g/mL from the mean. The computed density should then be expressed as:

$$1.298 \pm 0.017 \text{ g/mL}$$

The percent deviation of the density is computed to be $100 \times 0.017/\mathbf{1.298} = 1.3$ %.

This is close to the sum of the percent deviations of the mass and volume (0.1% and 1.3% respectively). This reflects a **general rule**, namely, that *the **percent** deviation in a quantity computed as the **quotient** of two quantities will be the sum of the **percent** deviations of the two individual quantities.* The same rule applies to quantities computed as the **product** of two quantities.

What about the deviations of computed quantities involving the *sum* or *difference* between two uncertain numbers? Consider the measurement of net volumes delivered by a buret. Suppose the initial buret reading is 0.74 mL and the final reading is 35.87 mL. Suppose that, due to the intrinsic precision of the buret, each of these readings has an uncertainty of 0.02 mL. The net volume is:

$$V_{net} = (35.87 \pm 0.02) - (0.74 \pm 0.02) = 35.13 \text{ mL}$$

What uncertainty should we ascribe to the net volume? Again, we look at the worst possible cases. The maximum net volume would result from using the largest final volume (35.87 + 0.02 = 35.89) and the smallest initial volume (0.74 − 0.02 = 0.72), i.e., 35.89 − 0.72 = 35.17 mL, and the minimum would be 35.85 − 0.76 = 35.09 mL.

These values each differ from the earlier result (35.13 mL) by 0.04 mL producing an average deviation of 0.04. The net volume should, therefore, be reported as:

$$V_{net} = 35.13 \pm 0.04 \text{ mL}$$

Note that the average deviation of the result is the sum of the average deviations of the two quantities. What about the percent deviation? The percent deviation of the initial buret reading is $100 \times 0.02/0.74 = 2.7\%$. The percent deviation of the final reading is $100 \times 0.02/35.87 = 0.06\%$. Because the final volume is large compared to the initial volume but the uncertainties are the same, its percent deviation is much smaller. The percent deviation of the computed net volume is $100 \times 0.04/35.13 = 0.1\%$. Note that the percent deviation of the result is much closer to the percent deviation of the quantity with the smaller percent deviation (in this case, the final volume). This again reflects a **general rule**, namely, *the **average deviation** of the sum or difference of two numbers is the sum of the **average deviations** of each of them.*

Types of Errors

Two types of errors contribute to experimental uncertainty—**random errors** and **systematic errors**. **Random errors** are those over which the experimenter has little or no control. They may result from **intrinsic precision** of the measuring devices (e.g., the same object weighed on the same analytical balance can give weights that differ by 0.0002 g simply due to the balance mechanism). They may also arise from

the lack of ability to control an experimental condition (e.g., a sample of a material that is analyzed may have small variations in composition from one sample to the next because of the way it was produced or constructed).

Using averages of several samples will tend to cancel out **random errors**. This may cause an increase in the numerical value of the expected *precision* of a measurement, but ***will generally not affect the accuracy of the result***. The word **homogeneous** (and the corresponding property, **homogeneity**) describes the extent to which a material has uniform properties no matter which sample is used.

When you are given an unknown material to analyze or reagents to use in an exercise, considerable effort is expended to insure that the samples are homogeneous to minimize such errors. When you dissolve a known weight of a substance to make a solution of accurately known concentration, it is important that the solution be mixed until it is **homogeneous**. If it is not homogeneous, different samples of equal volumes of the solution may contain significantly different amounts of the dissolved substance. Substances which are not of uniform composition are called **heterogeneous**.

Systematic errors are errors that are introduced by the experimenter or the experimental procedure. Suppose a small amount is lost in the process of transferring a weighed amount of a solid material from a container to a flask. The transferred material is used to make a *homogeneous* solution of known concentration. The concentration of the solution will be uniformly incorrect because of the lost sample. The subsequent procedure may be carried out with high precision, but the result is guaranteed to be inaccurate.

A common problem in the introductory chemistry laboratory occurs in the use of stock solutions of a precisely known concentration. If the student uses a wet beaker to transfer the solution, the water in the beaker causes the concentration of the solution used to be less than the indicated concentration. The exercise may result in very precise results which are inaccurate because of the decreased concentration. Results of high precision but low accuracy, are an indication of the occurrence of a systematic error.

Decimal Digits, Significant Figures, and Experimental Measurements

The appropriate use of significant figures plays a major role in the laboratory. Grading standards for exercises and quizzes invariably include a substantial number of points that deal with the proper implementation of this concept. We encourage the use of electronic calculators for facilitating the computations associated with experimental numbers. Even inexpensive calculators are designed to deal with and display numbers having 10 or more digits. Everyday laboratory measurements rarely permit many of these digits to represent reality.

Through significant figures, we seek to limit the number of digits in a quantity to those that are justified based on its origin. Two features play a role in the determination of significant figures. If a numerical quantity is the result of a direct measurement, the number of **decimal** digits in that quantity is generally fixed by the measuring device.

If it is a computed quantity, the number of **significant** digits reflects the number of digits in the numbers from which it is computed.

For experimentally determined quantities, the number of significant figures is **not** simply related to the intrinsic precision of a device. The buret permits a volume to be read to ±0.02 mL (one-fifth of the smallest division on its scale). How many significant figures does such a quantity have? If the recorded volume is less than 1.00 mL, the number of significant figures in the volume is two (e.g., 0.57 mL). If the recorded volume is between 1.00 and 9.99 mL, the number of significant digits is three. If the volume is between 10.00 and 50.00 (the maximum capacity of the buret), the number of significant figures in the volume reading is four—the maximum number of significant digits that a 50 mL buret is capable of producing. **Buret readings must always be reported to *two decimal places*, regardless of the volume reading**. Similarly, the display of an analytical balance displays weights to the nearest 0.0001 g. The number of significant figures represented by this precision varies with the total weight of an object. Nevertheless, **weights determined on an analytical balance must always be reported to *four decimal places*, regardless of the weight.**

The number of significant figures in a computed quantity may be the same or different from those of the quantities from which it is computed. Consider a container which together with its contents weighs 35.2749 g. A sample of the contents is transferred out, producing a final weight of 35.2235 g. The weight of the sample is 0.0514 g. The six significant figures in the initial and final weights produce a number with only 3 significant figures (but still with four decimal digits). Suppose the sample weight is now used in the calculation of a concentration by dividing it by a volume of 20.32 mL. Despite the four significant figures in the volume, the resulting concentration should be reported to at most 3 significant figures, namely .00253 g/mL. **In general, the number of significant figures in a multiplication or division of quantities will be that of the quantity with the least number of significant figures.**

> We will pay great attention to the use of significant figures in the grading of laboratory reports.** They reflect an understanding of the measurement limitations of the devices used in the exercises. You are urged to master this concept early in the laboratory course. Most introductory chemistry texts treat this topic and provide exercises to permit you to evaluate your understanding of their use.

Logarithms, Exponentials, and Scientific Notation

The treatment of significant figures in computations involving exponentials, logarithms and conversions between them, deserves some special discussion. We are accustomed to using scientific notation for the representation of large or small numbers, e.g., Avogadro's number is almost always written in the form 6.022×10^{23} mol^{-1}

**The use of, and emphasis on, significant figures in introductory chemistry is a hotly debated issue. In this course, their proper use represents the student's understanding of the precision of the measuring devices used in the exercises.

166

to avoid having to write the large number of 0's which serve only to specify the location of the decimal point. In this notation, the number of significant figures becomes easy to determine. All of the digits preceding the "X" are significant—in this case, four digits. If we write it to almost the full precision to which it is known, namely 6.022141×10^{23}, the number has 7 significant figures. The power of ten is irrelevant to the number of significant figures contained in the representation.

Scientific notation, then, is a way to remove any ambiguity in significant figures in a quantity. It may, however, not be a convenient representation when we are asked to add a series of numbers like $3.442 \times 10^1 + 5.645 \times 10^{-2} + 1.739 \times 10^2 + 7.21$. This is particularly evident when these numbers are added on a typical calculator, the result of which is 215.58645—eight significant figures. Writing the numbers as normal decimals permits one to see instantly that at most the four leftmost digits are significant in the answer. (You should add the numbers to confirm this!)

While scientific notation is a precise way of presenting data with the proper number of significant figures, the burden of keeping track of them falls entirely on the person doing the computing.

Logarithms (and the exponentials to which they are related) pose a similar problem. Consider the relation between a pH and its associated $[H^+]$ ($= 10^{-pH}$). If a solution has a pH of 4.15, how many significant figures are justified in the computed $[H^+]$? This is best recognized by writing $[H^+] = 10^{-4.15} = 10^{(-5 + 0.85)} = 10^{-5} \times 10^{0.85} = 7.08 \times 10^{-5}$. The pH has 3 significant figures, but one of them is basically a decimal place holder. To see that at most two significant figures are justified in the final answer, consider the following, where we increase and decrease the **final digit** in the pH by 1 unit (i.e., to 4.16 and 4.14):

$$10^{-4.16} = 10^{-5} \times 10^{0.84} = 6.92 \times 10^{-5} \text{ and } 10^{-4.14} = 10^{-5} \times 10^{0.86} = 7.24 \times 10^{-5}$$

Changes of 1 unit in the third significant figure of the pH causes changes of more than one unit in the second decimal place of $[H^+]$ calculated from the pH. We can formulate a rule of thumb that: in the process of converting a logarithm to the number it represents, we will get one less significant figure in the result.

What happens if we take the logarithm of a number?

$$[H^+] = 3.35 \times 10^{-3} \qquad \log(3.35 \times 10^{-3}) = -3 + \log 3.35 = -3 + 0.525 = -2.475$$

As we did before, adding and subtracting 1 in the third significant figure of $[H^+]$ gives us:

$$\log(3.36 \times 10^{-3}) = -3 + 0.526 = -2.474 \text{ and } \log(3.34 \times 10^{-3}) = -3 + 0.524 = -2.476$$

The corresponding change in the logarithm is one in the fourth significant digit!

Summary

Every experimental measurement involves uncertainty. Uncertainty can be due to a measuring device or due to the way it is used. Uncertainty is measured by accuracy and precision. **Accuracy** measures the deviation from the true value. For qualitative determinations, accuracy often means "true" or "false." For quantitative determinations, accuracy measures the deviation of a result from its "true" or accepted value. **Precision** measures the extent to which measurements that can be duplicated agree with one another. For small numbers of repetitions, the average deviation is an acceptable measure of precision.

When we calculate derived results from quantities with known precision, the precision of the calculated values depends on which arithmetic operations are involved. For addition and subtraction, the **average deviation** of the result is the sum of the **average deviations** of the quantities. For multiplication and division, the **percent deviation** of the result is the sum of the **percent deviations** of the quantities.

The number of decimal digits in a **measured** quantity is normally determined by the measuring device. The number of significant digits in a **computed** quantity depends both on the significant digits in the numbers from which it is computed and on what arithmetic operations are involved.

168

[Pre-laboratory Questions]

Name

Date   Course/Section

1. A student reports the following data in several determinations of the density of a homogeneous solid plastic sphere.

	RUN 1	RUN 2	RUN 3	RUN 4
Mass of sphere (g)	8.30	8.34	8.25	8.32
Volume of sphere (mL)	6.3	6.4	6.3	6.6

a. How should the student report the mass of the sphere? What is the percent deviation?

b. How should the student report the volume of the sphere? What is the percent deviation?

c. How should the student report the density of the sphere? What is the percent deviation?

d. If the intrinsic error of the mass determination is 0.03 g and the intrinsic error of the volume determination is 0.1 mL, what is the expected intrinsic error in the density?

e. If the true density of the sphere is 1.26, what are the deviation and percent deviation of the accuracy of the density determination compared with the true density?

(over)

170

f. Comment on how random or systematic errors might have been involved in this determination.

g. The grading standard for this exercise was as follows. What grade should be assigned?

	< 2% Error	2%–4% Error	4%–6% Error	> 6% Error
Accuracy	50 points	40 points	30 points	25 points

	< 3%	3%–4%	4%–5%	> 5%
Precision	50 points	40 points	30points	25 points

2. What are the average, average deviation, and percent deviation of the digits of your student ID?

3. How many significant figures are in each of the following quantities?

QUANTITY	NUMBER OF SIGNIFICANT FIGURES
0.003441	
1.003441	
3441.00	
6.023060	
3.441×10^{-3}	
1.2345090	
3.441×10^{3}	
10.4566	
2	
200	