# Evaluating Experimental Results

#### Martin Veillette

## The Essential Final Analysis

Even after you’ve completed all the measurements for your experiment, an equally important phase still remains: evaluating what your results actually mean. This evaluation phase transforms raw data into meaningful scientific conclusions.

Your primary objective in conducting an experiment is to make substantive statements about relationships between physical systems and theoretical models. You must clearly identify what statements you want to make and ensure they are accurate, comprehensive, and fully supported by your measurements. The specific evaluation approach depends on your experimental circumstances—whether you worked with a theoretical model, whether your measurements were dominated by statistical fluctuations, and other contextual factors.

## Approaching Evaluation with the Right Mindset

Before diving into specific evaluation techniques, two essential principles should guide your approach:

First, recognize that experimental results are precious resources. Whether they come from a multi-million dollar research program or a simple classroom exercise, your results represent unique, sometimes irreplaceable information. Honor this by extracting every possible insight from your observations and ensuring your final conclusions are as complete as possible.

Second, maintain unwavering objectivity. It’s nearly impossible to approach an experiment without some preconceptions about what “should” happen. However, you must discipline yourself to assess results objectively. If outcomes differ from your expectations or hopes, report them honestly and use them constructively to guide future investigations.

In academic settings, students sometimes misunderstand experimental objectives, believing the goal is to reproduce known values. If your measurement of gravitational acceleration yields 9.60 m/s² instead of the textbook 9.81 m/s², you haven’t failed—you’ve simply produced a measurement with its own characteristics and uncertainties. Rather than fretting over differences from established values, focus on making your own measurements as reliable as possible and accurately assessing their uncertainties.

When you measure quantities with established values, resist comparing your results until your analysis is complete. This builds confidence in your experimental process—confidence you’ll need when eventually measuring previously unmeasured quantities during professional research.

If your gravitational acceleration measurement is 9.60 ± 0.30 m/s², recognize that the uncertainty is as significant as the central value. The 3% uncertainty might reflect equipment limitations rather than experimental shortcomings. Textbook values often appear without context about the sophisticated methods and equipment used to obtain them. Your aim should be honest, objective reporting of results with appropriate uncertainty limits—not perfect reproduction of established values in limited laboratory time.

## The Evaluation Process: Four Essential Stages

The complete evaluation process consists of four key stages:

1. Calculate the values and uncertainties of basic measured quantities
2. Assess correspondence between experimental system and theoretical model
3. Calculate values of the properties you set out to measure
4. Estimate the overall precision of your experiment

Let’s examine each stage in detail.

### Stage 1: Calculating Elementary Quantities

Your first task is determining the values and uncertainties of the fundamental quantities involved in your experiment. Consider a pendulum experiment designed to determine gravitational acceleration. You’ll likely have measurements of pendulum length (ℓ) as your input variable and time measurements for multiple oscillations as your output, from which you’ll calculate oscillation period (T).

Your approach depends on whether you’re working with estimated uncertainties or statistical treatment of random fluctuations.

#### Working with Estimated Uncertainties

For pendulum length measurements using a meter stick, you’ve likely determined intervals within which you’re confident your true values lie. Your results would appear as a set of values in the form: value ± uncertainty.

Similarly, if you’ve counted oscillations and timed them with a stopwatch, you might express time measurements with their uncertainty ranges. However, the oscillation period T (your actual variable of interest) must be calculated from these measurements. If you counted 15 oscillations that took 18.4 ± 0.2 seconds, the period for a single oscillation would be:

(1/15)(18.4 ± 0.2) = 1.227 ± 0.013 seconds

Notice that both the central value and uncertainty must be calculated through this division. This significant modification of uncertainty values is necessary whenever you perform arithmetic operations on basic measurements.

Your final result will be a set of ℓ and T values with their associated uncertainties, preparing you for graphical analysis.

#### Working with Statistical Uncertainties

If repeated measurements show random fluctuations, you may have collected multiple readings for statistical analysis. You’ll need to express these as central values with uncertainties suitable for plotting.

As discussed in earlier chapters, sample means and standard deviations of means provide readily interpretable statistical significance. When reporting results, clearly indicate that you’re quoting sample means and standard deviations so readers understand you’re specifying intervals with 68% probability of containing the true value.

Remember that laboratory samples are often too small for definitive assessment of the underlying distribution. You’re making an assumption when applying Gaussian distribution properties to your sample, though it’s usually reasonable.

Also recall warnings about small-sample statistics—generally, statistical approaches aren’t worthwhile with fewer than 10 observations.

Consider how uncertainty regions will be interpreted on your graph. If both variables have similar statistical character, each point’s uncertainty rectangle will have clear interpretation. If variables have different uncertainty types (estimated versus statistical), interpretation becomes problematic. You might need to standardize them—perhaps using twice the standard deviation of the mean (95% probability) to make them comparable to estimated uncertainties.

At this stage, every experimental quantity should have a central value and uncertainty, but you’re not quite ready for graphing. If you need to plot derived variables (like T² vs. ℓ for a pendulum), you must calculate these through arithmetic operations. Remember to properly propagate uncertainties—if plotting T² values, uncertainty bars must represent the actual interval over which T² is uncertain.

### Stage 2: Creating Effective Graphs

Whether your graph serves as a simple illustration or as the key analytical tool, your goal is displaying results so their characteristics are immediately apparent. This requires thoughtful choices about scale, proportions, and presentation.

First, ensure your graph paper is sufficiently large. Plotting high-precision observations (0.1%) on standard letter-sized paper is futile when graphing uncertainty is around 2%. Unless uncertainties are clearly visible, you’ll lose valuable information. Similarly, make your graph fill the available area by choosing appropriate scales and suppressing zero when necessary. When plotting copper wire resistance versus temperature with values ranging from 57-62 ohms, starting your resistance scale at 55 rather than zero creates a meaningful display instead of a “flat roof” over empty graph paper.

There are exceptions where preserving the origin is important—when examining behavior near zero or when illustrating variation relative to baseline values. Generally, however, maximize use of graph space.

Clearly indicate uncertainty ranges for each measurement. You might use crosses with horizontal and vertical bars showing uncertainty ranges, or small rectangles encompassing the measurement with dimensions indicating coordinate uncertainties. The specific method matters less than consistently marking uncertainties on every graph. Note the nature of these uncertainties (estimated limits, standard deviations, etc.) on the graph itself or in its caption to prevent readers from hunting through text for interpretation guidance.

If plotting multiple datasets on one graph, differentiate them clearly through different symbols, colors, or other distinguishing features.

### Stage 3: Comparing Models with Experimental Data

Once your observations are plotted, you’re ready for the crucial step—comparing system properties with model predictions. The approach varies by circumstance, but we’ll assume you’ve arranged variables for linear graphical representation.

#### Scenario 1: Fully Specified Model

If you’re working with a completely specified model without undetermined parameters, your goal is simply assessing how well model predictions match experimental observations. Draw the model’s function on the same graph as your experimental points, using identical scales. This approach was illustrated earlier with falling object observations compared to the theoretical expression:

(1)

How do you judge correspondence quality? This is where uncertainty intervals become crucial. Without them, the inevitable scatter in experimental points would make meaningful comparison impossible—what are the chances of a theoretical line passing exactly through multiple scattered points? When points represent possible value intervals rather than single values, logical assessment becomes possible.

If the line representing the model passes through each point’s uncertainty region (as in the earlier example), you can state this observation directly. This doesn’t “prove” the equation is “true” or “correct”—it merely indicates the model and system are “consistent,” “in agreement,” or “compatible” within your measurement precision. Using appropriate language prevents misrepresentation and potential misunderstanding.

Also recognize that agreement exists only at your current precision level. At higher precision, discrepancies might appear that weren’t detectable in your experiment.

#### Scenario 2: Partial Correspondence

Sometimes a model adequately describes a system only within certain parameter ranges. Your graphical comparison might resemble Figure 6.1(b) or 6.1(c), showing agreement over limited ranges. For instance, fluid flow through a pipe might follow a linear pressure-flow relationship only below turbulence onset, or metal resistivity might follow a linear temperature model except at very low temperatures.

In such cases, report your comparison using language like: “We observed agreement between model and observations only over the range X to Y” or “The properties diverged significantly beyond value Z.”

Resist thinking something is “wrong” when models and systems don’t correspond completely. Both exist independently, and we cannot prejudge their overlap extent. Detecting validity limits for particular models often provides valuable clues for model improvement.

#### Scenario 3: Unexpected Intercepts

You’ll frequently encounter situations where a model’s behavior passes through the origin, but experimental observations don’t, as shown in Figures 6.1(d) and 6.1(e). Such discrepancies can arise from various model-system mismatches and provide valuable analytical insights.

When drawing graphs, check behavior at the origin. As previously discussed, graphical analysis helps obtain answers free from systematic errors associated with unexpected intercepts. Knowing whether such intercepts exist helps assess overall correspondence between model and system.

#### Scenario 4: Unexpected Data Scatter

During experiment planning, you should have carefully assessed measurement uncertainties and chosen appropriate methods to achieve your target precision. If you’ve done this properly, the scatter in your plotted points should be consistent with your estimated measurement uncertainties, as in Figure 6.1(a).

However, reality often deviates from expectations, and you may find yourself facing a situation like Figure 6.1(f), where scatter exceeds predicted uncertainty. This usually indicates unforeseen factors in your measurement process that weren’t accounted for in your initial uncertainty assessment.

Don’t leave such discrepancies unaddressed. Check your apparatus to identify potential fluctuation sources—perhaps a loose electrical connection or unstirred heating bath. Resolving such issues is always satisfying. If continuing the experiment isn’t possible, work with your existing results and make the best assessment you can of correspondence between model and system, perhaps noting that observations distribute uniformly around the model line.

#### Scenario 5: Complete Non-correspondence

It’s rare to encounter situations where system behavior bears no resemblance whatsoever to model behavior [Figure 6.1(g)]. With properly functioning equipment, this outcome is highly unlikely. Models may be imperfect representations of physical reality, but they wouldn’t qualify as models if they performed as poorly as this scenario suggests.

Such complete correspondence failure usually indicates experimental error—misinterpreting variables, incorrectly transforming equations, improper equipment setup, or mistakes in observation, calculation, or graphing. If possible, review everything from the beginning. If equipment access isn’t possible, check all analytical and arithmetic processes. If all error-finding attempts fail, report your results honestly and objectively. You may have discovered something novel, and an honest account of puzzling results from well-checked equipment will interest others in your field.

Throughout this assessment process, remember: experiments don’t give “right” or “wrong” results. Your responsibility is conducting experiments carefully and reporting outcomes honestly and objectively. Occasional reminders that models provide only partially satisfactory representations of physical systems are healthy. Understanding model validity limits and failure modes provides invaluable evidence for those seeking to improve them.

### Stage 4: Determining Values from Straight-Line Analysis

In previous sections, we discussed comparing fully specified models (including all numerical values) with experimental systems. However, as explained in earlier chapters, straight-line analysis frequently serves to determine unknown model parameters appropriate for your system.

In these cases, the model contains initially unknown quantities, so you cannot draw a complete model graph for comparison with experimental points. Your graph initially contains only the points themselves, as shown in Figure 6.2(a).

Consider measuring current through and potential difference across a resistor to test Ohm’s Law (V = IR). Without knowing resistance R, the model behavior encompasses all lines through the origin on the I-V plane described by:

(2)

where the constant could be any positive value. In principle, you could draw all possible lines on your graph and determine: (1) the extent to which system and model behaviors overlap, and (2) the range of R values appropriate for your system (as illustrated in Figure 4.11).

In practice, this is complicated by the fact that, based on measurements shown in Figure 6.2(a), you cannot assume system behavior passes through the origin. It’s best to defer the intercept question and simply determine which straight lines are consistent with your observations.

#### Finding the “Best” Line and Uncertainty Range

Several approaches exist for line-fitting. The most rigorous statistical method (least squares) will be discussed later. Meanwhile, we’ll examine simpler mechanical procedures, starting with the time-honored practice of drawing the “best” straight line through points by eye.

This requires mechanical aids that don’t obscure half your data points. Avoid opaque rulers; use transparent straight edges or, better yet, dark thread that can be stretched across points and easily repositioned. If visually judging point trends is difficult, hold the graph at eye level and sight along the points—this makes clustering around a straight line or systematic deviations much more apparent than direct viewing.

Identify several significant lines: the “best” straight line by your judgment, plus the limiting lines representing how far you can reasonably rotate your “best” line before it no longer acceptably fits the data. These extremes provide uncertainty values for your slope.

If wide point scatter makes identifying best-fit and limiting lines difficult, remember that your measured points represent samples from a continuous distribution band. The sparse population of this band (due to limited observations) can complicate line selection. Visualize the band populated by millions of potential readings your apparatus might produce, then estimate the center and edges of that distribution, allowing you to select appropriate lines.

In Figure 6.2(b), you might choose AB as your “best” line and determine that lines CD and EF would contain almost all possible points from an infinite measurement set. Lines CF and ED (not shown) would represent the steepest and shallowest slopes consistent with your observations.

Once you’ve selected appropriate lines, determine their slopes numerically to calculate your desired parameter (like resistance R in our Ohm’s Law example). For slope calculation, angle is irrelevant—you need the quantitative relationship between measured variables. For a line like AB in Figure 6.3, identify precise coordinates where it crosses graph grid intersections near its endpoints. If these coordinates are (I₁, V₁) and (I₂, V₂), calculate:

(3)

For our example, R equals this slope directly. In more complex cases, you might need additional calculations involving other measured quantities to determine your final answer.

Perform this process three times: once for your “best” line (AB) and once each for your upper and lower limiting lines (CF and ED). This gives your best value for R plus upper and lower limits beyond which you’re “almost certain” the true value doesn’t lie. Typically, these extreme values are roughly equidistant from your central value, allowing you to express your result as:

(4)

Sometimes your “best” line and limiting lines won’t appear equally spaced, usually because too few points prevent good line assessment. While sometimes experimenters feel compelled to express asymmetric uncertainties as:

(5)

visual graph judgment rarely justifies such precision. If identifying a clear “best” line proves genuinely difficult, you can simply delineate the edges of the value band (lines ED and CF in Figure 6.3), calculate maximum and minimum slopes, and express your experimental result as the interval between these slopes, or as their average ± half their difference.

If your desired answer isn’t directly equal to the slope, but requires calculation using additional quantities with their own uncertainties, combine the slope uncertainty with these other uncertainties using methods described in Chapter 2.

The significance of uncertainty values obtained from graphs depends on how you marked uncertainty on your original data points. If your bars represented outer limits of possible variation (either subjectively assessed or 2Sₘ for statistical fluctuations), your slope limits have similar interpretation. If points were marked with 1Sₘ limits, the limiting slopes probably represent better than 68% probability because of the conservative approach used in drawing limiting lines.

This analysis assumes that actual data scatter falls within predicted uncertainty ranges. If scatter greatly exceeds expected uncertainty (due to unforeseen fluctuation sources), you may have difficulty establishing lines that contain “almost all” possible values with confidence. In such cases and for all precision work, least squares analysis (discussed later) becomes essential.

When selecting your three lines, deliberately exclude the origin from consideration, as system behavior at the origin may be one aspect you wish to examine. If your model should pass through the origin, check whether the area between your limiting lines includes the origin. If so, your model and system show consistency at your precision level. Only if both limiting lines clearly intersect an axis on the same side of origin can you confidently identify an unexpected intercept.

If your model predicts an intercept from which you hope to determine some quantity, the intersection of your three lines with the relevant axis directly provides that intercept as: value ± uncertainty.

## Handling Imperfect Model-System Correspondence

When model and system correspond only partially, exercise care to avoid introducing systematic errors from these discrepancies into your results. Consider first cases where measurements align with the model’s straight line only over limited ranges [Figures 6.1(b) and 6.1(c)].

Obviously, restrict your slope evaluations to regions where system and model are compatible. Points systematically deviating from the straight line reflect physical circumstances not included in your model, making them inappropriate for model-based calculations. Disregard all points deviating systematically from straight-line behavior by amounts clearly exceeding estimated uncertainties and observed scatter, limiting your slope and uncertainty calculations to the linear region.

A second consideration involves intercepts. Even when your model passes through the origin, graphs frequently show intercepts. Such deviations arise from various causes, but many prove harmless. If the intercept-causing discrepancy affects all readings equally (like undetected zero error in an instrument or constant spurious EMF in an electrical circuit), the graph’s slope remains free from the systematic error that would otherwise contaminate your results.

Therefore, design experiments so answers come from graph slopes, while quantities potentially subject to undetermined systematic errors appear as intercepts. This capability to provide answers free from many systematic error types represents one of graphical analysis’s principal advantages.

## The Principle of Least Squares

All previously described procedures share a common limitation—they rely on experimenter judgment. While these approaches are useful and common, they invite criticism that even when carefully applied, their numerical significance remains uncertain. Having a mathematical procedure to identify the “best” line for a dataset would free us from judgment-related insecurity and potentially provide insight into what “best” actually means while allowing more precise uncertainty calculation.

The method meeting these needs is based on the statistical principle of least squares. We’ll focus primarily on its application to straight-line fitting, though it can be extended to other functions.

Consider a set of N (x,y) measurement pairs where uncertainty is confined to the y-dimension—we’ll assume x values are exactly known or sufficiently more precise than y values that x-dimension uncertainty can be neglected. This assumption is reasonable for many experimental situations where one variable is significantly more precise than the other. If both variables have comparable uncertainty, more complex treatments are needed (see Wilson’s text in the Bibliography).

Our mathematical procedure must answer: Which line on the x-y plane is “best,” and what does “best” mean? Least squares makes this determination based on vertical deviations of points from a candidate line. For line AB in Figure 6.4, consider vertical intervals between points and line (like P₁Q₁ and P₂Q₂). The “best” line minimizes the sum of squares of these deviations.

This criterion offers no automatic path to “truth” or “correct” answers—it’s simply one optimization criterion among many possibilities (we could minimize third powers or first powers of intervals, etc.). However, it can be proven that minimizing squared deviations produces smaller variance in resulting parameters (like slope) upon repeated sampling than any alternative criterion. This provides greater confidence in least squares results than competing approaches, explaining its near-universal adoption.

Mathematically, we define the best line as that which minimizes:

(6)

giving parameters (slope m and intercept b) for that line.

If our line equation is y = mx + b, each deviation δyᵢ equals the difference between measured y value and the corresponding point on the line:

(7)

The least squares criterion seeks to minimize:

(8)

with conditions:

(9)

Solving these equations (full derivation in Appendix A2) yields formulas for the best-fit line parameters:

(10)

(11)

We’ve now replaced potentially questionable visual judgment with a mathematical procedure yielding results of well-defined significance and universal acceptability. Since this method has statistical foundations, we can expect more precise uncertainty calculations. The least squares principle immediately provides standard deviations for slope and intercept, giving uncertainties with known statistical significance.

These standard deviations are calculated using the standard deviation of y-value deviations from the best line, Sy:

Sy = √[∑(δyᵢ)²/(N-2)]

Don’t worry about the N-2 denominator rather than the familiar N or N-1; it results from applying standard deviation definition to line positioning on a plane. The standard deviations for slope and intercept are:

(12)

(13)

Full derivations appear in Appendix A2.

These standard deviations, combined with m and b values, determine intervals with normal statistical interpretation—one standard deviation gives 68% probability of containing the true value, two standard deviations 95%, etc. A key least squares advantage is providing statistically significant uncertainty values for slope and intercept. These values derive objectively from actual point scatter, independent of any optimistic claims about measurement precision.

Appendix A2 also describes an extension for unequally precise data points, allowing greater weight for more precise measurements. This “weighting” procedure applies whenever we combine observations of unequal precision, even for simple tasks like finding the mean of unequal-precision values. Weighted mean and weighted least-squares calculation formulas appear in Appendix A2.

## Least-Squares Fitting for Nonlinear Functions

The procedures used for determining best-fit straight line parameters can, in principle, apply to nonlinear functions. We can write analogous deviation equations for any function and use similar minimization requirements for parameters in our chosen model. If resulting equations are solvable, we can find parameter values as we did for straight lines.

Frequently, however, these equations resist straightforward solution. In such cases, we abandon analytical approaches in favor of iterative computer solutions. We construct trial functions, calculate squared-difference sums, and progressively vary function parameters until finding the minimum sum. Computer-based methods for this process are described in Draper and Smith’s text (Bibliography). When possible, testing models in linear form remains simpler.

In all cases, experimenters are responsible for choosing appropriate functions—least squares merely determines which parameter values within a chosen function class best fit the observations.

## Important Cautions When Using Least Squares

Least squares mathematical procedures are entirely objective and impartial. Equations for linear fitting will drive a straight line through any dataset, regardless of whether a straight-line function is appropriate. If your experiment produces observations clearly showing breakdown of a linear model (Figure 6.5), blindly applying least squares to all observations will yield parameters for line AB that have no significance for either model or system. Thoughtless application of least squares methods must be scrupulously avoided.

This warning is particularly important given easy access to calculators and computers that generate least squares parameters for any inputted numbers with a few button presses. Remember that we compare straight lines with observations because we’ve judged this comparison reasonable. Therefore, never use least squares procedures before plotting observations and visually confirming linear fitting’s appropriateness. As mentioned earlier, you may need to exclude observations outside the model’s scope from best-line determination.

Only after carefully considering the entire situation graphically and visually, and confirming linear fitting’s appropriateness over all or part of the observation range, are you justified in applying least squares. Ignoring this warning can cause serious experiment interpretation errors.

## Finding Functions When No Model Exists

Our previous discussion assumed you possessed a model to compare with your system. While this is common, sometimes you encounter observation sets with no available model—perhaps when researching previously unobserved phenomena or studying systems too complex for theoretical modeling. When plotted, such observations typically show curves with no readily identifiable pattern. Without a model, what approaches are available?

One option is finding functions with some correspondence to your observations. This can be valuable in complex systems where theoretical modeling seems hopeless. Even if your “model” is merely a mathematical function restating the system’s behavior, it facilitates computer processing and enables interpolation, extrapolation, and similar operations. Such empirical models help predict national economic responses to taxation changes or determine temperatures from resistance thermometer calibration curves.

In simpler systems where theoretical modeling seems possible, functions showing good correspondence with observations may guide model building by suggesting underlying physical processes. However, caution is essential. Finding a function consistent with observations at a particular precision level doesn’t “prove” you’ve discovered the “right” function. Different function types often show similar behavior over limited variable ranges, and guidance from incorrectly identified functions can be misleading, potentially impeding theoretical progress for years. The physics history contains many examples where researchers failed to recognize empirical function choices must remain provisional.

With appropriate caution regarding potential limitations, here are some common function-finding approaches:

### Power Law Functions

As discussed in experiment planning chapters, logarithmic plotting helps identify power law relationships. Consider the function:

y = xᵃ

Taking logarithms of both sides:

log y = a·log x

A graph of log y versus log x produces a straight line with slope a. To test whether observations follow a power law, plot them as log y versus log x. If points align with a straight line, you can conclude a simple power function (positive/negative, integral/fractional as determined by the graph) fits your observations well. The appropriate power value comes from the graph’s slope, with uncertainty limits depending on plotted point uncertainties.

Such graphs can use ordinary paper (plotting actual log x and log y values) or logarithmic graph paper (with rulings proportional to logarithms, allowing direct plotting of original values).

### Exponential Functions

Many physical phenomena follow exponential relationships:

y = ae^(bx)

Taking natural logarithms:

ln y = ln a + bx

This creates a straight line when plotting ln y versus x (a “semi-log plot”). If you suspect an exponential function might apply to your system, create a semi-log plot using either ordinary graph paper (calculating ln y values) or semi-log paper (with one logarithmic and one linear scale). Appropriate a and b values come from the intercept and slope, with uncertainties determined by measurement precision.

### Polynomial Approximations

If neither power laws nor exponential functions adequately match your observations, the chances of finding a more complex function that fits well are slim. In such cases, polynomial approximations often prove useful:

y = a₀ + a₁x + a₂x² + ...

While such representations essentially admit ignorance about underlying system mechanisms, they still offer empirical modeling advantages—facilitating computation and providing bases for interpolation and extrapolation.

Finding appropriate coefficients for such expansions typically employs the least squares principle. As noted earlier, computational difficulty increases rapidly with the number of terms needed for satisfactory correspondence. Fuller discussion appears in Draper and Smith’s text (Bibliography).

Similar approaches apply when observation scatter isn’t severe and highest precision isn’t essential. Finite difference calculus techniques can be applied to observations, and difference tables used for interpolation, extrapolation, or polynomial fitting. Comprehensive discussion appears in texts by Whittaker and Robinson and by Hornbeck (Bibliography), with elementary treatment in Appendix A3.

## Assessing Overall Experimental Precision

At experiment initiation, you estimated likely uncertainties to guide your experimental approach. After completion, you should retrospectively evaluate actually achieved precision through critical results assessment. The specific uncertainty type matters less than clearly stating what you’re reporting—whether estimated ranges, standard deviations, standard deviations of means, or other measures.

For meaningful application, overall uncertainty figures must be realistic and honest, even when experimental outcomes are less favorable than hoped. Include all identifiable uncertainty sources in your assessment. If balance points cannot be identified within 2-3 mm or slide wire non-uniformities introduce errors, claiming 0.2% precision for slide-wire potentiometer readings becomes meaningless, regardless of millimeter-graduated scales.

Known systematic error contributions should be excluded at this stage, as appropriate measurement corrections should already have been applied. However, suspected systematic error sources whose contributions cannot be accurately evaluated should be described with appropriate allowances in overall uncertainty ranges. Final statement format depends on circumstances:

### For Results Based on Measurement Sets

The best quantity to report is standard deviation of the mean, which has recognizable numerical significance. Sometimes standard deviation itself is quoted. Always specify sample size so σ estimate reliability can be judged.

### For Results from Single Calculations

If graphical analysis wasn’t possible and results come algebraically from several measured quantities, use Chapter 3 methods to calculate either outer uncertainty limits or standard deviations.

### For Results from Graphical Analysis

If your straight line was established through least squares, constant uncertainties (m and b) will have been directly determined. These uncertainties advantageously derive from actual point scatter, independent of estimated uncertainties. (This doesn’t mean you should skip plotting uncertainties or drawing graphs when using least squares—as emphasized earlier, graphs with plotted uncertainties remain essential for judging model-system correspondence ranges before least squares calculations).

If you’ve drawn your line by eye, the limiting possibility lines will give slope and intercept ranges. This slope uncertainty may need combining with other quantity uncertainties before stating final answer uncertainty.

As mentioned previously, the specific uncertainty type matters less than clearly stating what you’re reporting. When working through lengthy uncertainty calculations, you can simplify by dropping insignificant contributions—adding 0.01% uncertainty to 5% offers negligible benefit since the 5% value lacks three-digit precision. Final uncertainty statements rarely justify more than two significant figures; only highly significant statistical work warrants greater precision.

Once you’ve determined overall answer uncertainty, consider how many significant figures to retain. This was covered in Section 2.11, but bears repeating in the context of experiment evaluation.

No unique answer exists for significant figure questions, but generally avoid retaining figures beyond the first uncertain digit. For example, 5.4387 ± 0.2 should be reported as 5.4 ± 0.2, since uncertainty in the tenths position makes subsequent digits meaningless. If uncertainty is known more precisely, retaining one additional figure might be justified—if uncertainty were 0.15, reporting 5.44 ± 0.15 would be valid.

When reporting percentage precision, significant figures are automatically implied. A measurement reported as 527.64182 ± 1% implies absolute uncertainty of 5.2764. However, since precision is quoted to just one significant figure (1%, not 1.000%), the uncertainty itself warrants only one significant figure. Calling it 5 implies the tens digit in the original number is uncertain by 5, making subsequent digits meaningless. The measurement should be quoted as 528 ± 5 or 528 ± 1%.

For sample means, significant figures depend on the mean’s standard deviation, which in turn depends on the standard deviation’s standard deviation.

Finally, always ensure answer and uncertainty expressions are consistent—neither “16.2485 ± 0.5” nor “4.3 ± 0.0002” represents good practice.

## Understanding Correlation

Thus far, we’ve considered experimental interpretation involving relatively precise observations and satisfactory models. Reality is often messier, and much modern experimentation is less clear-cut than previous sections might suggest.

Many scientific fields deal with subtle phenomena where effects can be partially or completely masked by statistical fluctuations or other perturbations. In these scenarios, detailed model-system comparisons may be impossible—you might struggle even to establish whether the effect you’re studying exists at all. This scenario commonly occurs in biological, medical, and environmental studies.

Consider familiar public health debates about smoking’s role in lung cancer, low-level radiation’s relationship to leukemia, or dietary influences on cardiovascular disease. In these contexts, “proof” frequently enters discussion: “We haven’t proved smoking causes lung cancer” or “Can we prove heart attacks are less frequent with margarine versus butter consumption?”

These scenarios operate in fundamentally different domains from our earlier experimental approaches. Understanding what we mean by terms like “proof” and “cause” becomes critical.

Consider two experimental scenarios. First, measuring current through a resistor as potential difference varies, producing results like Figure 6.6(a). Have we “proved” current is “caused” by potential difference? The current certainly differs between low and high potential differences by amounts far exceeding measurement uncertainty, giving confidence the variation is real. Was this variation “caused” by potential difference changes? We observed current increases with potential difference increases, but theoretically, current might be unrelated to potential difference, with increases caused by some entirely separate factor like atmospheric pressure, making the apparent relationship purely coincidental.

Philosophers have warned for centuries that simultaneous events aren’t necessarily causally related. However, accumulated experience with this experiment, involving multiple repetitions and careful control of other variables, gradually convinces us potential difference and current are genuinely related. Only philosophical purists would dispute that potential difference causes current flow.

The situation differs dramatically in less clear-cut cases. Another experiment might yield results like Figure 6.6(b), typical when studying, for instance, university student cold incidence versus daily vitamin C consumption. Can we conclude cold frequency depends on vitamin C dosage? We might conduct a well-designed experiment with 100 students receiving vitamin