# Designing Experiments: Principles and Methods

#### Martin Veillette

## Introduction to Experimental Design

In the previous chapter, we explored the various ways researchers compare models with real-world systems. The diversity we encountered suggests a crucial insight: there is no universal approach to planning experiments. The techniques and methodologies researchers employ necessarily depend on specific circumstances and objectives.

Despite this diversity, certain fundamental principles remain valid across virtually all experimental situations. Perhaps most important among these is keeping your experimental purpose clearly in mind: **the fundamental requirement in experimentation, regardless of what else is happening, is to compare the properties of a physical system with the properties of one or more theoretical models**.

Let’s assume that preliminary investigation has already identified the significant variables in your system. Some variables will be under your control (input variables), while others will be determined by the system itself (output variables). We’ll primarily focus on situations where input variables can be isolated and controlled independently. When everything varies simultaneously, interpreting results becomes significantly more challenging—a common obstacle in professional research, but one we’ll avoid for now by focusing on fully controlled experimental scenarios.

## Testing an Existing Model

When you already have a theoretical model—whether it’s a simple relationship like Hooke’s Law (F = kx) or Ohm’s Law (V = IR), or something derived from a sophisticated theory like Einstein’s general relativity—the model’s properties will typically take the form of functional relationships between variables.

Your primary goal remains comparing the model’s properties with those of the physical system. Only after verifying through experimentation that the system and model properties overlap adequately (at least within some range) can you confidently proceed with measuring the quantities of interest.

Remember that determining whether a model is appropriate for a given system must be based on experimental evidence. We aren’t attempting to decide whether models are “true” or “false” in some absolute sense—all models are imperfect approximations. Rather, we need to determine if a particular model is adequate for our specific purposes at our desired level of precision.

If conditions change or greater precision becomes necessary, we must reconsider the model’s adequacy. As discussed previously, graphical approaches typically provide the most effective way to test physical models. Ideally, we want to plot the model’s behavior alongside our experimental observations of the system’s behavior, which requires some preparation.

Since conventional graphs are two-dimensional, we initially need to limit ourselves to examining relationships between two variables at a time. When dealing with multiple input variables, we can simplify by holding all but one constant while studying how the output variable depends on the remaining input variable. After completing this analysis, we can adjust one of the previously fixed variables and repeat the process. Through successive measurements of this kind, we can construct a comprehensive picture of the system’s behavior.

This approach assumes we can hold input variables constant independently of one another. When this isn’t possible, more sophisticated techniques become necessary, which we’ll touch on later.

For now, assuming we’re working with a single input variable (either because only one exists or because we’ve isolated one by controlling the others), our procedure is straightforward: measure how the output variable changes with the input variable, then plot these measurements for comparison with the model’s predictions. As noted earlier, the advantages of linear representation are so significant that we’ll focus primarily on transforming our data into straight-line form.

## Converting Equations to Straight-Line Form

### Basic Transformations

When a model contains only linear functions—such as distance traveled at constant velocity versus time, or potential difference across a constant resistor versus current—the equation is already in straight-line form. However, this simplicity is rare. More commonly, we need to transform the functions in our model into linear form.

Consider a function describing the time of fall for an object:

(1)

To represent this in linear form:

(2)

We might choose:

(3)

(4)

(5)

(6)

There’s no single formula for these transformations. The most effective approach is keeping the target form clearly in mind while rearranging the original equation until you achieve the desired structure.

Multiple valid transformations often exist for a given equation. The function above could be equivalently expressed as:

(7)

(8)

(9)

While convention often suggests plotting input variables horizontally and output variables vertically, there’s no strict requirement to do so. Choose the representation that best serves your analytical purposes.

### Practical Considerations

Your plotting choices should consider both experimental requirements and practical convenience. Generally, plot variables in their simplest form, leaving complex arithmetic for the final calculation stage.

For example, when determining viscosity using fluid flow through a pipe, Poiseuille’s equation gives us:

(10)

Where:

* Q = flow rate (volume/time)
* P = pressure difference
* a = pipe radius
* ℓ = pipe length
* η = viscosity coefficient

You might consider plotting Q versus (πa⁴/8ℓ)P to get a slope of 1/η, but this introduces unnecessary complications. Each pressure measurement would require multiplication by πa⁴/8ℓ, and uncertainties in a and ℓ would artificially inflate the uncertainty in your compound quantity.

A better approach: plot Q versus P directly, using πa⁴/8ηℓ as the slope. This simplifies both your plotting process and uncertainty analysis. After determining the slope, calculate η using:

(11)

This principle—plot variables in their simplest form and leave arithmetic for the final calculation—will serve you well in experimental design.

### Working with Compound Variables

Sometimes convenience or necessity requires plotting variables constructed from multiple primary measurements. Consider a compound pendulum—a rigid body oscillating about an axis perpendicular to its plane. For small oscillations, its period is given by:

(12)

Where:

* T = oscillation period (output)
* h = distance from center of mass to support point (input)
* g = gravitational acceleration (unknown constant)
* k = radius of gyration about center of mass (unknown constant)

Converting this to linear form using single-variable functions of h and T proves impossible. However, using compound variables makes it possible. Starting by squaring both sides:

(13)

Multiplying both sides by h:

(14)

Rearranging to isolate h²:

(15)

This gives us a linear equation where:

* Vertical variable = h²
* Horizontal variable = T²h
* Slope = g/4π²
* Intercept = -k²

This example demonstrates the superiority of linear analysis. The traditional approach—plotting T versus h—provides only k’s value (from intercepts) and requires calculating g indirectly. Linear analysis offers numerous advantages:

1. It provides a clear basis for comparing the system with the model
2. It allows reliable uncertainty estimation
3. It avoids the unreliability of shallow-angle intercepts
4. It utilizes all data points rather than just those near intercepts
5. It determines g and k independently, preventing error propagation

Compound variables also prove valuable with multiple input variables. When measuring specific heat using flow calorimetry, the heat balance equation is:

(16)

Where Q is heat generation rate, m is mass flow rate, C is specific heat, and ΔT is temperature difference.

Rather than plotting ΔT versus 1/m (with separate curves for different Q values) or ΔT versus Q (with separate curves for different m values), you could plot mΔT versus Q. This creates a single graph incorporating both input variables simultaneously, with slope C, enabling efficient model testing and parameter determination.

If plotting with compound variables reveals unexpected patterns (scattered data or nonlinearity), you can always revert to plotting individual variable pairs to investigate further.

### Logarithmic Transformations

Many physical processes involve exponential relationships:

(17)

Taking natural logarithms of both sides:

(18)

Plotting ln y versus x (a “semi-log plot”) produces a straight line with slope b and y-intercept ln a. While base-10 logarithms would also work, only the intercept value would change.

Logarithmic plotting applies to simple power relationships too. For:

(19)

Taking logarithms:

(20)

Plotting log y versus log x (a “log-log plot”) yields a straight line with slope n.

This approach offers two key advantages over plotting y versus xⁿ directly:

1. It accommodates very large variable ranges on reasonably sized graph paper
2. It reveals when data follows a power law with an unexpected exponent

For instance, if measurements follow y = x¹·⁸ rather than y = x², plotting y versus x² would show systematic deviation from linearity without revealing the true relationship. A log-log plot would still produce a straight line, immediately indicating a power relationship, with the slope revealing the actual exponent (1.8).

We’ll explore log-log plotting further when discussing empirical model construction in the next chapter.

## Step-by-Step Experimental Planning

Let’s examine the practical steps for preparing an experiment. While these may seem excessive for simple teaching laboratory exercises, they represent essential practices for serious research. The simple experiments in educational settings simulate more complex situations you’ll encounter in professional work, where inadequate planning can have serious consequences.

The planning process includes:

1. **Identify system and model**: This seemingly obvious step can be surprisingly challenging. The phenomenon under study is often surrounded by measurement apparatus, obscuring the fundamental system. If you struggle to identify your system, ask: “What entity’s properties does my model describe?”

Similarly, clearly define your model’s limitations. When studying falling objects, will you account for air resistance? Neglecting air resistance isn’t irresponsible—it’s defining one aspect of your model. The experiment itself will reveal whether this simplification is justified at your desired precision level.

1. **Select variables**: Typically, one quantity presents itself as the natural output variable. If there’s only one input variable, selection is straightforward. With multiple input variables, identify your primary independent variable and vary others in discrete steps.
2. **Transform the equation**: Put your model equation into straight-line form as described earlier. Remember, multiple valid transformations usually exist. Choose one that serves your purposes effectively. When the equation contains unknown parameters to be determined experimentally, structure your transformation to place these unknowns in the slope rather than the intercept whenever possible. Intercepts are more susceptible to systematic errors from instrument defects.
3. **Determine variable ranges**: Plan for an input variable range spanning at least a factor of 10. Wider ranges provide better basis for comparing system and model behaviors. While you can’t directly control output variable ranges, carefully consider instrument limitations. Circuit components have maximum current ratings, materials have elastic limits, and sensors have operating ranges. Perform trial measurements to determine input variable ranges that avoid damaging equipment or exceeding measurement capabilities.
4. **Consider experimental precision**: Begin with a target precision level for your final result. This guides your measurement method selection. A request to “measure g using a pendulum” is meaningless without specifying whether you need 10% precision (achievable with simple equipment in minutes) or 0.01% precision (requiring sophisticated apparatus and days of work).

With a clear precision goal—say, measuring g within 2%—you can work backward to determine requirements for each component measurement. For a pendulum experiment, if you need g within 2%, you might aim for uncertainties in length (ℓ) and period-squared (T²) below 1% each.

If you can measure length with ±1mm uncertainty, the minimum acceptable length measurement would be:

(21)

(22)

Similarly, if timing uncertainty is ±0.2s, and period measurement requires 0.5% precision (for 1% in T²), the minimum timing interval would be:

(23)

(24)

This analysis helps ensure all measurements contribute meaningfully to your desired final precision. If any measurement appears limited to uncertainties exceeding your target, you’ll need either more precise measurement methods or a revised precision goal.

After completing these planning steps, create a comprehensive measurement program—typically a table listing all quantities to be measured and providing space for calculations needed in graphing. This allows you to focus on conducting the experiment without constantly deciding what to measure next, and helps prevent accidentally omitting crucial measurements.

The complete experiment design process is illustrated in Appendix A4 with a sample experiment.

While this planning may seem excessive for simple laboratory exercises, it represents the minimum preparation required for serious research. Resist the temptation to rush into measurements and figure out analysis later—developing good planning habits now will serve you well throughout your scientific career.

## Designing Experiments Without Existing Models

When investigating phenomena so new that theoretical models don’t exist, or systems so complex that constructing theoretical models seems impossible (like national economies or biological systems), your experimental approach necessarily differs.

Without an existing model to test, your motivations might include:

* Simple curiosity
* Practical information needs
* Seeking guidance for theoretical model construction
* Developing an empirical model for calculation purposes

Even without detailed theoretical understanding, empirical models prove extremely valuable. They help organize thinking about complex systems and enable mathematical operations like interpolation, extrapolation, and forecasting.

In model-free situations, experiment design becomes more straightforward if input variables can be isolated. Simply measure the output variable across suitable ranges of input variables to construct a comprehensive picture of system behavior. When input variables can’t be isolated, we encounter more complex challenges discussed in a later section.

Even without established theories, consider any available hints about potentially applicable functions, testing them against your observations. One powerful technique for obtaining such hints is dimensional analysis.

## Dimensional Analysis

Even without a complete theory, dimensional analysis can provide valuable guidance by examining the physical dimensions of quantities involved in your experiment. Physical quantities are expressed in terms of basic dimensions: mass (M), length (L), and time (T).

The fundamental principle: dimensional consistency must exist between both sides of any physical equation. For instance, if gravitational acceleration (g) relates to pendulum length (ℓ) and period (T), dimensional analysis reveals that g (with dimensions LT⁻²) must incorporate length to first power (L¹) and period squared (T⁻²):

(25)

This approach can’t determine dimensionless constants (like π), but it reveals functional relationships between variables.

The general method:

1. Express the relationship as a proportionality with unknown exponents: z ∝ xᵃyᵇ
2. Write dimensions of both sides using M, L, and T
3. Create equations by matching powers of M, L, and T
4. Solve for the unknown exponents

For example, analyzing the velocity (v) of waves on a string under tension (T) with mass per unit length (m):

(26)

Dimensionally:

* v: LT⁻¹
* T (tension): MLT⁻²
* m: ML⁻¹

Therefore:

(27)

Matching powers:

* For M: 0 = a+b
* For L: 1 = a-b
* For T: -1 = -2a

Solving gives a=½, b=-½, yielding:

(28)

This powerful technique becomes more complex with many variables. When analyzing fluid flow rate (Q) through a tube with pressure difference (P), radius (r), length (ℓ), and viscosity (η), we get:

(29)

We can determine some relationships (Q contains P/η and r³) but not all (the exponent b remains undetermined). Even this partial solution provides valuable experimental guidance.

## Difference-Type Measurements

Sometimes clear relationships between variables aren’t readily apparent, or input variables can’t be easily isolated. The system might be influenced by many factors, making it difficult to detect the specific effect you’re studying.

### Null-Effect Measurements in Physical Sciences

When studying subtle effects potentially masked by external factors (like measuring tiny extensions in steel wire affected by temperature fluctuations), use **null-effect** measurements. Study two identical specimens simultaneously—one exposed to your variable of interest, one not—and measure the difference between them.

For steel wires, you might cut a single wire in half, load one piece while leaving the other unloaded, and measure the difference in length. Both experience identical temperature variations, but only one responds to loading. This approach reveals small effects that would otherwise be lost in environmental noise.

Always check system performance both with and without the influence you’re studying. Wilson’s humorous observation is worth remembering: “It has been conclusively proved by numerous tests that the beating of drums and gongs during a solar eclipse causes the sun’s brightness to return.”

### Control Groups in Biological Sciences

Biological systems present additional challenges—you can’t simply “cut a specimen in half” when testing drug effectiveness. Human subjects show inherent variability that would overwhelm the effects you’re studying if comparing individuals.

The solution: compensate with numbers. Create an **experimental group** receiving the treatment and a **control group** that’s as similar as possible but doesn’t receive the treatment. Both groups experience the same external influences, but only one receives the intervention you’re studying.

This approach often requires refinements like placebo treatments and double-blinding (keeping both experimenters and subjects unaware of group assignments) to prevent psychological effects from contaminating results.

Such experimental designs—pairing treatment groups with carefully matched control groups—are foundational in biological research, whether studying carcinogenic food additives in mice or music education effects on academic performance.

## Observational Studies with Uncontrollable Variables

Sometimes you must study systems offering no experimental control whatsoever. Astronomers can’t manipulate celestial bodies, public health researchers can’t randomly assign people to “smoking” and “non-smoking” groups, and environmental scientists can’t create control planets.

In such cases, careful observational techniques become crucial. With well-defined systems and models (like celestial mechanics), precise measurements may still allow meaningful conclusions—determining that general relativity better explains Mercury’s orbit than Newtonian mechanics, for instance.

More challenging questions arise in complex systems: Does a manufacturing process change affect product quality? Do fluoridated water supplies improve dental health? Do nuclear power plants increase leukemia rates nearby?

These questions present multiple challenges:

* Little/no control over input variables
* Wide variation in individual responses
* Probabilistic rather than deterministic effects
* Delayed responses to interventions
* Inability to observe true null effects (we rarely have baseline measurements before interventions)
* Numerous confounding factors

Your best approach: meticulous sampling procedures. Create artificial null-effect measurements by constructing treatment groups under the influence you’re studying and control groups exempt from it but otherwise matched as closely as possible.

The validity of such studies depends entirely on sampling quality. Effects are often subtle enough that different sampling approaches can yield contradictory conclusions. This reality explains why public policy debates featuring scientific components often present competing “scientific” evidence—different sampling approaches can support opposing conclusions.

When facing such complexity, conventional concepts like “proof” require modification. Mathematical theorems can be proven from axioms, and some physical measurements are certain enough to be considered “proven” (the moon is closer than the sun). But in complex systems with probabilistic effects, “proof” gives way to **correlation**—statistical relationships between variables that differ fundamentally from direct cause-effect relationships but remain valid for identifying influencing factors.

We’ll examine correlation analysis further in Chapter 6 when discussing experimental evaluation.

## Problems

## Practice Problems

1. A research paper claims the terminal velocity of a skydiver depends solely on the skydiver’s mass and gravitational acceleration. Evaluate the reasonableness of designing an experiment to test this hypothesis.
2. A projectile is launched with initial velocity v at angle α to the horizontal. Its range may depend on the projectile’s mass, initial velocity, launch angle, and gravitational acceleration. Determine the functional form of this relationship through dimensional analysis.
3. The internal pressure in a soap bubble depends on the liquid’s surface tension and the bubble’s radius. Using dimensional analysis, determine the relationship between these variables.
4. A torsional oscillator’s period depends on the support’s torsional stiffness coefficient (torque per unit angular displacement) and the moment of inertia of the oscillating object. Find the functional relationship between these quantities.
5. The central deflection of a beam with circular cross-section, supported at both ends and loaded at its center, depends on the applied force, distance between supports, beam radius, and the material’s Young’s modulus. Use dimensional analysis to determine the relationship.

*For problems 6-23, state which variables or combinations of variables should be plotted to verify the proposed relationship, and explain how to determine the unknown parameter(s) from your graph (slope, intercept, etc.).*

1. The position of an object under constant acceleration follows:

(30)

where s and t are measurable. Determine the acceleration a.

1. The fundamental vibration frequency of a stretched string is given by:

(31)

where f, ℓ, and T can be measured. Determine m.

1. The exit velocity of an ideal fluid flowing through an opening in a tank follows:

(32)

where v and P are measurable. Determine fluid density ρ.

1. A conical pendulum’s period is described by:

(33)

where T and α are measurable, and ℓ is fixed and known. Determine g.

1. The deflection of a cantilever beam follows:

(34)

where d, W, and ℓ are measurable, while a and b are fixed, known values. Determine Young’s modulus Y.

1. The height of capillary rise in a tube follows:

(35)

where h and R are measurable, and ρ and g are known constants. Determine surface tension σ.

1. The ideal gas law states:

(36)

where p and T are measurable, and V is fixed and known. Determine gas constant R.

1. The Doppler frequency shift for a moving source follows:

(37)

where f and v₀ are measurable quantities, and f₀ is a known constant. Determine velocity v.

1. Thermal expansion of a solid follows:

(38)

where ℓ and ΔT are measurable, and ℓ₀ is unknown but constant. Determine coefficient of expansion α.

1. Snell’s law of refraction states:

(39)

where θ₁ and θ₂ are measurable angles, and n₁ is a known constant. Determine refractive index n₂.

1. The thin lens equation states:

(40)

where s and s’ are measurable. Determine focal length f. Compare two possible plotting methods and explain which is preferable.

1. The resonant frequency of a parallel LC circuit follows:

(41)

where ω and C are measurable. Determine inductance L.

1. Coulomb’s law for electrostatic force is:

(42)

where F and r are measurable, while q₁ and q₂ are fixed, known values. Describe how to verify the form of this relationship.

1. The force between parallel current-carrying conductors follows:

(43)

where F, i₁, i₂, and r are measurable quantities, while μ₀ and ℓ are constants. Describe how to verify this relationship.

1. The charge remaining on a discharging capacitor follows:

(44)

where Q and t are measurable, and R is known and fixed. Determine capacitance C.

1. The impedance of a series RC circuit follows:

(45)

where Z and ω are measurable. Determine resistance R and capacitance C.

1. The relativistic mass variation with velocity follows:

(46)

where m and v are measurable. Determine rest mass m₀ and speed of light c.

1. The wavelengths in the Balmer series of hydrogen follow:

(47)

where λ and n are measurable. Determine Rydberg constant R.