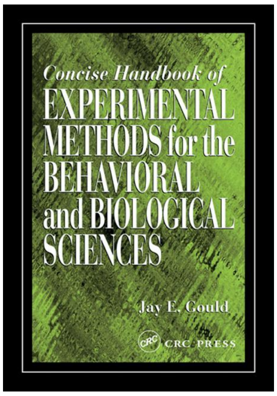

Concise Handbook of Experimental Methods for the Behavioral and Biological Sciences

Jay E. Gould

Steps of the Scientific Method

Publication details
https://www.routledgehandbooks.com/doi/10.1201/9781420040869.ch3
Jay E. Gould
Published online on: 20 Dec 2001

3

Steps of the Scientific Method

CONTENTS

Introduction

This chapter begins with a brief coverage of all the *primary and secondary steps of the scientific method*, and then examines Primary Steps 1 and 2 in detail. The other primary steps are elaborated later: Primary Step 3 in Chapters 4, 9, and 10; Steps 4, 5, and 6 in Chapters 11 and 12; and Step 7 in Chapter 5. With regard to the additional steps: Secondary Step 1 is elaborated in Chapter 11, whereas Secondary Steps 2 and 3 were discussed earlier in Chapter 1. The following are some important *general points*:

1. There is a good deal of *overlap and interrelatedness* among the steps.
2. Scientists, therefore, often *move back and forth* between these steps.
3. The steps might be *listed somewhat differently* by different scientists.
   — Such variations, however, just represent *complementary approaches* to the general scientific method.
   - Example*: Inductive method versus deductive method* are two important complementary strategies used in science (covered later in this chapter under Hypothesis Formulation).

Primary Steps

There are *seven major steps* involved in the scientific method of research: 1) formulate a problem, 2) formulate a hypothesis, 3) design a study, 4) collect and organize the data, 5) summarize and statistically analyze the data, 6) evaluate the results and draw conclusions regarding the hypothesis, and then 7) communicate the findings.

1. **Formulate a Problem**
   — The *starting point* of any scientific inquiry is a *research problem/question* about the nature/description of, or the explanation for, some phenomenon.
   - Example: What is the *cause* of clinical depression?
2. **Formulate a Hypothesis**
   — Next, a *hypothesis* is put forth as the possible *answer/solution* to the problem's question.
   - Typically, this is a statement about the *conjectured relationship* among the factors that will be investigated (the independent and dependent variables).
     - ∞ Example: A *deficiency* in the serotonin neurotransmitter mechanisms of the brain is the cause of clinical depression.
3. **Design a Study**
   — The *design* is a *plan* for gathering data to *test the hypothesis*.
   - Scientists must determine whether the hypothesis is *probably true or probably false*.
     - ∞ In other words, does the hypothesis appear to *solve the problem* — does it *answer the research question*?
   - *Experimental designs* are the most powerful designs for research (discussed earlier in Chapter 2).
4. **Collect and Organize the Data**
   — *Empirical observations/data* are the primary criteria for testing hypotheses (and theories); however, *thought experiments* — involving *rationalism* as an approach to knowledge — are also sometimes used.
   - Example: *Einstein* used thought experiments in working out and evaluating (i.e., "testing") his *Theories of Relativity.*
5. **Summarize and Statistically Analyze the Data**
   — *Descriptive statistics* are used to *summarize* the data.
   - Examples: *Measures of central tendency* — mean, median, and mode; *measures of variability* — standard deviation and variance; and *measures of relationship/association* — correlation coefficient (covered in Chapter 11, under "Statistics Versus Parameters").

   — *Inferential statistics* are used to *analyze* the data.

- Examples: chi-square test, *t*-test, and analysis of variance.
  - ∞ Such *tests of significance* are used to determine the *reliability*, not the magnitude, of the findings; specifically, to get a measure of the *probability* that the *association(s)* found between variables are only due to *chance* (tests of reliability versus magnitude are covered in Chapter 12, under "Statistical Significance Versus Practical Significance").

6. **Evaluate the Results and Draw Conclusions Regarding the Hypothesis**

— This step involves the processes of *interpretation* and *inference*.

- Data are *compared* to the hypothesis, which is a *prediction* of the results, to determine whether the hypothesis is or is not *supported* by the empirical evidence — and hence whether the hypothesis is *probably true or probably false*.
  - ∞ In experiments this can lead to *explanation*, i.e., specification of *causal relationships* among variables/phenomena.

7. **Communicate the Findings**

— Normally this last step involves the writing and publication of a *research report*.

— Instead or additionally, communication can involve an *oral or poster presentation* at a convention/conference.

— Communication increases the *utilization and impact* of the results.

— It also allows *independent* evaluation and verification of findings.

- This possibility is a *requirement* of scientific observation, i.e., the observations must be *public*.

— Although communication represents the *final major step* of the scientific method, it also might represent a *beginning*.

- Research reports serve as an important source of *new research problems and hypotheses* in the *ongoing process* of science.

Secondary Steps

In addition to the seven primary steps of the scientific method, there are *three closely related ancillary steps*: 1) generalization, 2) prediction, and 3) replication.

1. **Generalization**

— This secondary step is *closely allied* with the primary step of *drawing conclusions*.

— Scientists want to *extend* a study's findings from the *specific* set of *conditions* and *participants* used, i.e., they want to make as *general* a statement as possible about the *implications* of their results.

- However, they must be careful not to generalize *too broadly* since this could lead to *error*.

∞ Example: Generalizing research results obtained with a *sample* of U.S. college students to the *population* of all humans on the planet could be a mistake, since the sample might not be *representative* of the diverse world population (see Chapter 11 for an extended discussion).

2. **Prediction**

— Predicting is *closely related* to the secondary step of *generalization*.

- Generalizations, in fact, are actually a form of *prediction*.

— Scientific predictions often refer to the *proposed results* of a test of some prior *hypothesis/theory*, or a modification of it, under *new conditions* and/or involving *different types of participants/subjects*.

- Example: Proposing that the research results found for *rats* under *laboratory conditions* would also be found for the human species in our *complex environments.*

— This brings the scientist back near the beginning of the scientific method.

- Note: A significant piece of research, through generalization and creative thinking, will not only *generate new questions/problems and predictions/hypotheses*, but it will provide *theoretical implications* relevant for explaining *other phenomena* occurring in analogous situations.

3. **Replication**

— Repeating a study involves a special form of *prediction*.

- In this case, a study is *repeated* as it was *originally* conducted, in order to see if the results can be *reproduced*, and thus *confirmed/verified* — which would be the logical *prediction*.

— *Generalization* is typically involved to some extent since, due to practical limitations, *replication is rarely complete or precise*.

— *Partial replication*, rather than "complete" replication, is said to occur when something *other than* just the participants, experimenter, or location is *changed*, and/or when anything is *added or deleted*.

Problem Formulation (Step One of the Scientific Method)[26]

Developing a meaningful research problem is the first step in conducting scientific research (*meaningfulness* is discussed under items 5 and 6 below).

1. **Best Way to State a Research Problem**

— Usually it is *clearest* when a problem is put in the form of a *question*.

- Generally, research questions ask *how or why* something occurs, and specifically they ask *what the relationship* is between two or more variables — usually *independent and dependent variables*.

- However, in descriptive research the question might simply ask *what are the properties* (i.e., characteristics and their values) of some phenomenon.

2. Ways to Find a Research Problem

There are *three common approaches* to coming up with research problems to study:

a. Be alert to *personal experiences* and *newspaper or magazine reports*.

- *Everyday events* at work, school, home, athletic competitions, etc., can present practical issues that need solutions — so be *inquisitive*, develop a *questioning attitude*.
  - ∞ Example: What are the *motivations* for some unusual behavior(s) you have observed among fraternity members?

b. Read *journals and books*

- *Previous research and theories* raise many new questions that need answers.
  - ∞ Example: Given the *Theory of Evolution by Natural Selection,* how can *altruistic/helping behavior* — which might lead to the *death* of the helpful individual — be explained?

c. Attend *conferences*, take *courses*, and have *discussions*

- *Keeping current* in your knowledge and *tapping the minds of others* can lead to interesting ideas for research.

3. Ways in Which Research Problems Become Evident

There are *three types of situations* that make us aware of research problems that need to be addressed: a) contradictory results of studies, b) a gap in knowledge, or c) a fact that needs explaining.

a. Contradictory results of studies

- Two or more studies attempting to answer the *same question* might obtain *different results* and thus *different answers*.
- This naturally raises questions as to *why* this has happened.
  - ∞ Example: Investigations over a number of years have found opposite results for the effects of *heightened sexual arousal* on the level of *subsequent aggression* (sometimes finding that it increases it, and in other cases that it decreases it).
- Possible reasons for contradictions
  1) *Extraneous variables* might have been *inadequately controlled* in one or more experiments, thereby causing different results.
     - Probably the *most common cause* (see Chapters 7 and 8 for coverage of extraneous variable types and their experimental control).
     - As a result, one or more of the studies might have been *poorly conducted*, and thus the results would be questionable, or the *value* of one or more extraneous variables simply might have *varied* between the studies.
     - *Additional research* is usually needed to determine the cause of contradictions.

Example: In a study it was found that "Right handers *moved their eyes* leftward when solving spatial problems and rightward for verbal problems *when the questioner sat behind them*. But, *when facing the questioner*, the same participants moved their eyes predominantly in only one direction, either right or left, regardless of problem type. The *results indicate* that the cerebral hemispheres, though specialized for problem type, are also preferentially activated within an individual."

What is important for us, however, is that this study cleared up the inconsistent results appearing in the literature regarding direction of eye movement in response to problem type. The discrepant results were apparently due to an *extraneous variable* in the procedures used, specifically, *experimenter location*, which varied between the two studies. [Gur, E. R. (1975). Conjugate lateral eye movements as an index of hemispheric activation. *Journal of Personality and Social Psychology*, Vol. 31 (4), 751-757]

2) *Independent variable* values might have been different.
   - This could lead to different results in two or more studies if the *relationship* between the independent and dependent variables is *nonlinear*, rather than being linear. It occurs when the value of the dependent variable *does not change by a constant amount* as the independent variable is changed by a constant amount — thus a graph of the relationship would be *curvilinear*, rather than a straight line.

     This is particularly problematic when the relationship is *nonmonotonic*, i.e., when the *direction* of the independent variable effect *changes* for different portions of its range/continuum — in which case, as the independent variable increased in value the dependent variable value would first increase and then decrease in value, or vice versa (see figure in Chapter 7 under "Maximizing the Primary Variance").
3) *Dependent variables* might have been different.
   - When this occurs, the studies could be *measuring* different things, at least to some degree, or the studies could differ in their *sensitivity or range of possible values*.

     Example: Measuring the *speed* versus the *accuracy* of dependent variable responses is really very different.
4) *Participant populations* sampled might have been different.
   - Hence, they could differ in their *responses* to conditions.

     Example: Studying *college students* versus the *general population*.
5) *Designs* might have been different.
   - Some designs are *more powerful or sensitive* than others. (Examples are covered in Chapters 9 and 10.)

6) *Statistical analyses* might have been different.
   - Here too, some are *more powerful/sensitive* than others. (Examples are covered in Chapter 12.)
7) *Type I or Type II Errors* might have occurred.
   - These errors are covered in Chapter 12 under Potential Errors During Statistical Decision Making.

b. Gap in knowledge

- Being aware of *what we do not know,* as a result of what we do know, is another way that research problems become evident.
  - ∞ It is often said that: "The more we *learn*, the more we realize we *don't know.*"
  - ∞ Example: Knowing that *marijuana* has a debilitating effect on *short-term memory* raises questions about what other effects this drug might have on *cognition.*

c. Fact that needs explaining

- Being *aware* of a fact but not understanding *why it is so* is yet another way in which research problems become evident.
  - ∞ A fact needs explaining when it *does not fit in with* — i.e., cannot be *related to* — existing knowledge.
  - ∞ It *demands* explanation, and thus *collection of further data*.
- Problems of this kind might lead to *major discoveries*.
  - ∞ Example: Knowing that *opiates,* such as morphine, relieve pain and are very addicting led to a search for their mechanisms and the eventual mapping of the distribution of opiate receptors in the brain, and subsequently to the even more important discovery that the brain produces its own *endogenous opiates* that regulate pain, mood, and learning.

4. Basic Types of Problems — Classes of Questions[24,31]

There are four fundamental types of research problems. Distinctions made here between *proximate versus ultimate questions* help us to *avoid confusion* about the mechanisms/causes of behaviors and mental processes (cognition, emotion, and motivation) and help to *organize and guide research*. However, the proximate and ultimate mechanisms are *interrelated and complement one another* — they are not competing explanations, rather they are *different levels of explanation*.

a. Proximate causal mechanisms (two "how" questions)

1) Immediate causation (stimuli and mechanisms)
   - ∞ What are the *external environmental factors* (e.g., discriminative, sign, or releaser stimuli) and the *internal psychological, physiological, and anatomical factors* (e.g., goals, endocrine, and neural events) that are involved in the short-term regulation (i.e., elicitation, control, and execution) of the behavior or mental process?

2) Development (ontogeny)

∞ What are the *hereditary/genetic/innate factors* and the *environmental/experiential/learning factors* that are involved in assembling the behavior or mental process over longer time spans in the lives of individuals, and what are the relative roles? (Note: This relates to the classic Nature-versus-Nurture Problem.)

b. Ultimate or distal causal mechanisms (two "why" questions)

1) Adaptive significance (function)

∞ What is the current role of the behavior or mental process in *promoting survival and reproduction*, i.e., what are its advantages for gene replication — How is it adaptive?

2) Evolutionary history (phylogeny)

∞ What factors *influenced the evolution* — and what was the evolution — of the behavior or mental process during the species' history over geological time?

Example: What might the answers be to the four different types of questions above for the phenomenon of *human language*?

5. Restriction on Problems

— Meaningfulness

- To be meaningful, the question put forth by the problem *must be capable of being answered* with the tools *available* to the *scientist*.

  ∞ Thus a *meaningful* problem is a *solvable* problem.

- Present meaningfulness

  ∞ Research problems with present meaningfulness can be solved with the tools *presently available* to the scientist.

  - Only this type of problem/question can be considered for *current* research.

- Potential meaningfulness

  ∞ Research problems with only potential meaningfulness are *not presently meaningful*, but it is considered *possible* that *in the future* the tools necessary to solve them will become available to the scientist (see below).

  - This type of problem/question should, therefore, be placed in a *wait-and-see category*.

6. Meaningless or Presently Meaningless Types of Problems

There are *four reasons* for problems *not* being presently meaningful: 1) there is no empirical reference, 2) it is presently impossible to obtain relevant data, 3) the problem is unstructured and vaguely stated, or 4) the terms are inadequately defined.

a. No empirical reference

- For such problems it is impossible to obtain *relevant objective data* in order to solve them, hence they are *meaningless*.

- ∞ These problems do not deal with properties *observable* by the ordinary senses of human beings, even using instruments.
  - Example: "What is the *mind of God* like?"
- ∞ Hence such problems are *not* subject to scientific investigation.

- These problems *only can be solved* by divine vision, revealed truth, or similar mystical power; by intuition; or by reasoning, i.e., *rationalism*.
  - ∞ Such problems, therefore, are described as being not scientific, but rather *metaphysical, mystical, theological, or superempirical*.

b. Presently impossible to obtain relevant data

- Although not presently solvable, such problems *could have an empirical reference*, and thus be at least *potentially meaningful* for the future, after advances in research capabilities.
  - ∞ Example: "What is the biochemistry of life forms *outside our solar system*, if there are any, and what rules govern the behavior of these life forms?" Although we *cannot presently* gather the necessary data, advances in space travel or communication *might* allow this in *the future*.

c. Unstructured and vaguely stated

- In such problems, it is *not clear* what the *relevant variables/ events* are and thus what the *appropriate observations* would be.
  - ∞ Example: "How does the *universe* operate?"
  - ∞ Such questions are *too general*, and thus it is *uncertain* what is meant by them — making the problems *meaningless*.
  - ∞ Note, however, that it might be possible to *reformulate* the question so that it is *more precisely stated*, and thus *more specific*, and hence *answerable*.
    - Example: "How do stars form and how does *life* evolve?"
  - ∞ *Experimental problems* should be *clearly stated* in terms of specific, observable, *independent and dependent variables*.
    - Example: "What are the effects of a *stimulus-enriched environment* on the *structure and functioning of neurons* in the brain?"

d. Terms inadequately defined

- Poorly defined terms or variables are a source of *vagueness*, hence this cause for problems being meaninglessness is really a *subset* of the preceding reason.
  - ∞ Once again, it is unclear exactly what the necessary *relevant observations* would be.
- *Clear specification* is needed of the meaning of the problem's important terms if productive research is to be conducted.
- Major cause for the *ambiguity* of terms is that our everyday language is full of *words with multiple definitions*.
  - ∞ Examples: Without precise definitions, how can we study the possible causes of *love*, or the effects of *pornography?*

∞ To make concepts clear and unambiguous, science uses *operational definitions*, which give terms/variables an *empirical basis* (discussed later near the end of this chapter).

7. Other Considerations Regarding Research Problems

a. Interest

- Is the problem *intriguing enough* to merit consideration?
  - ∞ Problems are said to be *uninteresting* if the answer to the question put for by the problem is *obvious*.
    - Caution: The obvious is not always the *actual outcome* of carefully controlled research — *science has its surprises*.

      Example: In studies on the increase in the number of neurons that occurs during the development of the nervous system, who would ever have expected that *before birth half the neurons in the brain actually die?*

b. Importance

- Is solving the problem *worth the time and effort*?
  - ∞ Caution: Experiments on what *appears to be a trivial problem* could produce results that are *very important*.
    - Example: Studying the *behavior of bears and coyotes* has led to non-harmful ways of controlling the undesirable behaviors of harming humans or killing livestock. One approach developed is to eliminate open garbage dumps, and another is to produce conditioned taste aversion. The psychologist John Garcia has done much useful *basic and applied research* in this area.

c. Expense

- Problems might be quite important and interesting but *too costly*.
  - ∞ Example: Are there *other intelligent life forms* in the universe, and just what are their *intellectual abilities?*

d. Basic/pure research Versus applied research

- Are either basic or applied research problems *more crucial*?
  - ∞ Basic research
    - This category of research investigates problems relating to *fundamental questions* about the principles and mechanisms of various phenomena.

      Example: What are the anatomical and biochemical mechanisms in the brain of *learning and memory*?
  - ∞ Applied research
    - This category of research investigates problems whose solutions are expected to have *immediate application*.

      Example: What events in the brain are responsible for the memory loss occurring in *Alzheimer's disease*?

- Research on both types of problems is essential.
  - ∞ *Basic research* establishes the foundation for applied research.
  - ∞ *Applied research*, in turn, can improve the quality of life of humans and other species, and also raises new questions for basic research.
  - ∞ It would be *very shortsighted* for society to concentrate on applied research at the expense of basic research, since applied research is *dependent* on basic research.

e. Chance of contributing new knowledge

- Some problems are more likely to *advance science* than others.
  - ∞ All things being equal, the scientist should *focus* on those problems that are most likely to yield *productive outcomes*.
    - Example: When a problem develops from *studies yielding contradictory results,* to conduct *just another experiment* wouldn't seem worthwhile if the only outcome would be to *score another point* for one side in the controversy.

      To be of *real value*, the new study would have to use an *innovative, fresh approach*, one that would be likely *to explain* the contradictory results, and thus *more thoroughly* answer the research question.

      In other words, it is important to know what, exactly, were the *critical differences* in the earlier studies that *account for the contradictory findings.*

      More specifically, were different studies investigating the effects of different ranges of an *independent variable*, or measuring somewhat different *dependent variables*, or were *extraneous conditions* different?

Hypothesis Formulation (Step Two of the Scientific Method)[26]

Having developed a research problem that is expected to be solvable, and thus meaningful, the next step is to propose a *meaningful hypothesis*.

1. Definitions of Hypotheses and a Restriction

a. Hypothesis

- Broadly speaking, a hypothesis is a *potential answer* to the question put forth by a problem.
  - ∞ It is a *tentative solution* to the problem posed by a scientist.

b. Meaningful hypothesis

- A *meaningful* hypothesis is a *testable* hypothesis.

- ∞ It is a *testable proposition* that might be a problem's solution.
  - As was noted for problems, *scientific hypotheses* also should be *restricted* to those that are meaningful.
  - It must be possible, in other words, to determine a *degree of probability* (there are few absolutes) as to whether a hypothesis is true or false — and thereby to *reject or fail to reject* the hypothesis.
  - Note: A *problem* is considered meaningful (solvable) if there is a meaningful (testable) *hypothesis* that is relevant to it — hence this is a very useful way of *checking* on whether a *problem* is in fact *meaningful*.

c. Meaningful experimental hypothesis

- This type of hypothesis is an *experimentally testable* statement of a *potential relationship* between *independent and dependent variables* that might be the/a solution to a research problem.
  - ∞ It must be testable *empirically* through the *purposeful, direct manipulation* of antecedent conditions, and the *measurement* of consequent conditions, under *controlled* conditions.
  - ∞ Note that the definition of a meaningful *experimental* hypothesis incorporates the preceding definition of a *meaningful* (i.e., testable) hypothesis, which in turn incorporates the basic, general definition of a *hypothesis*.

2. Best Ways to Write Hypotheses

a. Logical form of the general implication

- This manner of writing a hypothesis is what is commonly called a *conditional statement*: "If ____, then ____."
  - ∞ Generally put: *If* certain *antecedent conditions* occur, i.e., some specified *manipulation* of an *independent variable(s)*; *then* certain *consequent conditions* will occur, i.e., some specified *measured value(s)* of the *dependent variable(s)*.
  - ∞ It is important to note that this represents a *prediction*.
    - To be *testable*, hypotheses must make *predictions* about *observable events*, not about unobservables.
  - ∞ Conditional statements are usually *clearer* and *more concise* than the typical, less structured ways in which hypotheses are *commonly stated* by students, as well as professionals.
    - If not initially phrased as a conditional statement, it is nonetheless *possible and advantageous to rephrase* such hypotheses in the "If ____, then ____" form.

      Example: Paying someone for work they usually do voluntarily will result in a reduction of the quality of the work that they do." This can be restated more clearly and concisely as: "*If* someone is paid for work they usually do voluntarily, *then* the quality of their work will be reduced."

Note that by using this *logical format* it becomes very clear what will be *done* and what is *expected*, i.e., what will be *manipulated* and what will *be measured*.

b. Quantitative form of a functional relationship

- Stating a hypothesis as a *mathematical function, or equation*, is very desirable, but often it is not yet possible to do so in the behavioral sciences, or sometimes even the biological sciences.
  - ∞ The general form of the *mathematical function* (*f*) would be

    dependent variable = *f*(independent variable or variables)

    - Example: Psychological perceptual responses (ψ), e.g. perceived loudness, are a *specific function* of physical stimulus values (l), e.g. sound intensity — such as was proposed in *Fechner's Log Law,* i.e., ψ = k(log l) + C (Note: k and C stand for constants that are computed).
  - ∞ Note: The quantitative method of writing a hypothesis is also a *conditional statement*, but it is a *more precise* "If ____, then ____" expression.
    - Rather than just stating, e.g., that "*if* the independent variable increases, *then* the dependent variable will increase;" it is stated that "*if* the independent variable is increased by *X* amount, *then* the dependent variable will correspondingly be increased by an amount equal to *Y*."

3. Principle of Parsimony

— According to this principle: If *different hypotheses* are presented to answer the question posed by a given problem, then the one that is *least wasteful of resources* — i.e., *most parsimonious* — is preferred.

- This is a *basic rule and assumption of scientific thinking*, which is usually right, about the organization/operation of the universe.

— Two components

- Occam's Razor
  - ∞ *General explanations* are to be preferred over those that are appropriate to a more limited range of phenomena, according to this component of parsimony.
    - Thus if two hypotheses (or theories) are equally complex, the one that can *explain more results* is to be preferred.
    - In other words, explanatory principles should not be *needlessly multiplied*.

      Hence it should be *assumed* that *one set of principles* explains a wide range of phenomena under a variety of conditions, until *empirical evidence* indicates that this assumption is incorrect.

This axiom is appealed to when *generalizing* research *results* and their *explanations* across different types of individuals and species, different independent and extraneous variable conditions, and different dependent variable measures.

- Morgan's Canon
  - ∞ *Simplest explanation* (hypothesis or theory), i.e., the one involving the *fewest explanatory concepts and assumptions*, is to be preferred when it accounts for the same amount of data just as well as do more complicated explanations.
    - Thus, this component is to never appeal to a *higher or more complex process* for explanation when a *lower or simpler one* (physically, psychologically, developmentally, evolutionarily, etc.) will do the job equally well.
  - ∞ Example of both Occam's Razor and Morgan's Canon: If mice are discovered in a barrel of flour when none were there before, we shouldn't jump to the explanation/theory of *spontaneous generation* of life from inanimate matter. *A more general and simpler explanation or hypothesis* is that there was a *small opening* in the barrel through which the mice were able to crawl.
- IN SUMMARY: *Seek the smallest possible number of simplest principles that can successfully explain the greatest number of phenomena studied*.
  - ∞ Note, however, that this *assumption* that "*less is more*" can sometimes lead to *error* — i.e., while usually correct, the principle of parsimony is *not always correct*.
    - As Albert Einstein (1879–1955) once said: "Everything should be made as simple as possible, but not simpler."

4. Theory

— Definition

- *Theories* are general *propositions* used to *explain* a phenomenon or, more commonly, a set of phenomena (empirical relationships).
  - ∞ Example: *Signal-Detection Theory* states that "Sensitivity to a stimulus depends not only upon the stimulus intensity but also upon the experience, expectation, and motivation of the person being tested."*[The Encyclopedic Dictionary of Psychology* (3rd ed.). *(1986).* Guilford, CT: Dushkin.]
- Hypothesis (idea, view, or notion) and theory are both terms referring to tentative explanations, and thus they are closely related and sometimes used interchangeably — nevertheless, technically these terms do have different meanings.
  - ∞ *Theories* are explanations of how *larger pieces of the puzzle* go together, as opposed to smaller pieces, i.e., they are usually explanations of *groups of phenomena* — thus theories are *more general* than hypotheses.

- ∞ *Theories* are also more likely to have *considerable supporting evidence* — thus theories are *less tentative*.

- *Principle*
  - ∞ This label/designation is typically used after a theory becomes more firmly established by the accumulation of *additional supporting evidence*.
    - Example: *The Phylogenetic Principle* that, in general, ontogeny (development of the individual) recapitulates phylogeny (evolutionary development of the species).
- *Law*
  - ∞ This label is commonly used after enough evidence has been obtained so that there is essentially *no doubt* about the veracity of some specified empirical relationship(s).
    - Example: *Law of Effect*, by E. L. Thorndike (1874–1949) and followed up by B. F. Skinner (1904–1990), that a response which is followed by a *pleasant consequence* is more likely to be repeated, and a response followed by an *unpleasant consequence* is less likely to be repeated.

— Functions

Theories have *two general purposes*. They are used to *organize established facts*, and they are used to *guide future research*. Hence theories look both to the *past* and to the *future*.

1) Organize established facts
   - ∞ Theories are an important aid for arriving at a *systematic body of knowledge*, and thus advance the other goals of science as well.
     - They *clarify knowledge* by providing a basis for the *organization of data* resulting from *many studies*.

       Example: It was theorized that there are separate systems for *short-term versus long-term memory*, based on such things as differences in storage capacity and rates of forgetting, as well as the effects of blows to the head, electro-convulsive shocks, and damage to brain areas.
     - Theories also *facilitate retrieval of knowledge* in a manner analogous to remembering a *rule* rather than all the *individual facts*.

       Example: *Drive-Reduction Theory*, which holds that motivated behavior arises from drives or needs, and that responses which satisfy these drives/needs tend to be reinforced, and thus are more likely to reoccur.
2) Guide future research
   - ∞ Theories also generate *testable predictions* by way of *specific hypotheses* derived from the theories, which can advance the scientific goals of *explanation and control* of phenomena.

- To be testable, a theory must make predictions about *observable events* (as also noted earlier for hypotheses).
- *Scientifically meaningful theories* are *testable theories*.

  Example: *Darwin's Theory of Evolution*, where the environment can be varied and effects observed on the survival and reproductive success of individuals as a result of changes in anatomical, physiological, or behavioral characteristics of successive generations.

— Meaningless theories

There are *two reasons* that theories might be meaningless, i.e., untestable: they might be *vague* or *too general*.

1) Theories are scientifically meaningless when stated in such *vague terms* that it is *unclear* what operations should be performed in order to *test* them — which is done by gathering *empirical evidence* that either supports or refutes the theory (consider the following: *immorality* is caused by *godlessness*).

   ∞ This point also applies to *hypotheses*, and is similar to the vagueness concern noted earlier for research *problems*.

2) Theories are also considered scientifically meaningless when they are *so general* that they can explain *any possible outcome*, and thus are *not refutable*, which means that they are *not testable*.

   ∞ Some theories are not refutable because their explanations are after the fact (*post hoc*), i.e., *no predictions are made*.

   - Example: *Freud's Personality Theory* involves the id, ego, and superego that vie for control. But the decision as to which is dominant at any time tends to be *after the fact*, thus the theory does *not* generate testable predictions.
   - It should be acknowledged, nevertheless, that a *scientifically meaningless theory* still can be *therapeutically meaningful*, i.e., useful. From the practitioners point of view, a theory is only as meaningful as the effective repertoire of *techniques* that it generates.

— Evaluation criteria

The *quality* of a theory is based on how *completely and accurately* it fulfills the *two* functions of theories noted earlier.

1) How well does the theory account for *past* research findings, i.e., how well does it *organize established facts*?

   ∞ It should be possible to generate *old results* from a theory.

   ∞ The *more* results that can be *explained*, and the *simpler* (i.e., fewer explanatory concepts) and more *precisely* this can be done, then the more *parsimonious* and *better* the theory.

2) How well does the theory predict *new* research findings, i.e., how well does it *guide future research*?

∞ It should be possible to generate *novel results* from a theory.

∞ The *more* results that can be *predicted*, and the *simpler* and more *precisely*, then the *better* the theory.

- Hence, *exact mathematical statements* are superior to more general verbal statements.

∞ *Strong tests of theories*, in contrast to weak tests, pit one theory against another.

- This is done by having them *generate different predictions* that are then compared against the data to determine which theory *best accounts* for the findings (see example later in this chapter under Deductive Method).
- This *comparative approach* is a *superior* means for evaluating the *relative quality* of theories or hypotheses.

— Additional considerations

There are *three other concerns*.

1) Mutual exclusiveness

∞ An important principle to always keep in mind is that different theories/hypotheses about the same phenomena are *not necessarily mutually exclusive*, i.e., if one is *correct*, it is not necessarily the case that the other is *incorrect*.

- When there is *good support for more than one theory*, the different theories might each explain *different aspects* of the same phenomena, or *different subsets* of the data — thus, both could be correct, at least in part.
- Example: *The Trichromatic Theory of Color Vision* (developed in the 19th century by Young and Helmholtz) and the *Opponent-Process Theory* (developed about the same time by Hering) were later both found to be correct, but for *different levels* in the visual system, i.e., peripheral retinal photoreceptors versus more central levels of the nervous system, respectively.

2) Mutual exhaustiveness

∞ Another important consideration — which is related and complementary to mutual exclusiveness — is whether or not proposed theories are *mutually exhaustive*, i.e., whether there might be *additional or better explanations*.

- Example: The *Place Theory of Pitch Perception* and the *Frequency (or Telephone) Theory* (developed in the 19th century by Helmholtz and Rutherford, respectively) were both found to be correct, but for coding *different ranges* of frequencies, i.e., relatively high frequencies versus low frequencies, respectively. However, later on, the *Volley Theory of Pitch Perception* was developed (in the 20th century by Wever and Bray). This theory *complements* the earlier two theories by *best accounting* for coding of the "middle" range of auditory frequencies.

3) Data versus theory

∞ Finally, a very important point is that *research data should always dominate theory* — not vice versa.

- If research results (empirical observations) do not support a theory, then the theory probably needs to be modified or even discarded — i.e., *the facts can't be changed to accommodate a theory (or hypothesis).*
- However, a fact of life is that a *bad theory* usually is *not eliminated* just because it is poor, but rather only because a *better theory* has been formulated that can take its place.

Example: Darwin's Theory of Evolution by *Natural Selection* was needed to replace Lamarck's Theory of Evolution by *Inheritance of Acquired Characteristics* (the necks of giraffes were said to have become longer with each generation because the young inherited the stretched necks of their parents who were reaching higher for leaves to eat).

5. Inductive versus Deductive Methods[33]

— General features

- The inductive and deductive methods represent *two variations* on the *general scientific* method.
- They are associated with the two different *functions of theories*:
  - ∞ *Induction* is used for *organizing/systematizing established facts*.
  - ∞ *Deduction* is used for *guiding future research* through the generation of predictions.
- The two methods actually represent *complementary strategies* for research, with the *weaknesses* of one being the *strengths* of the other (as discussed later).
- Moreover, although these different approaches proceed in *opposite directions*, they are *strongly interrelated*.
  - ∞ *Observations* (experimental or naturalistic) lead through *induction* to new *general principles or theories*, which in turn suggest through *deduction* what *further observations* should be made — thus coming full circle (this is elaborated later):

observations → *induction* ↓
↑ *deduction* ← theories

  - ∞ Investigators therefore do not necessarily use one approach to the exclusion of the other, although they might *prefer* one.

— Inductive method

- This approach is likely to be used when studying topics that have received little or no previous attention, i.e., *relatively unexplored areas* — thus little data would have been gathered, and hence good general principles or theories would not yet have been formulated.
- Induction moves from the *specific to the general*.
  - ∞ It begins with relatively *specific empirical observations or reflections*, which are then followed by the formulation of a *problem* (a question) and perhaps also a *hypothesis*.
  - ∞ Next it moves through the collection, organization, analysis, and summarization of a substantial *database*, obtained under varying conditions, onward to the eventual identification of *generalities* — the formulation of *general principles or theories*.
    - Example: Development of the theory that there are at least *two general forms of memory — procedural* (for how to do things) versus *declarative* (for facts and events) — based on observations of many amnesiac patients, which were followed up on by several additional studies.

      To elaborate: The observation was made that people suffering from amnesia typically remember *how* to do things, such as the use of language and driving a car, but they do not remember *who* they and others are, or *what* their past experiences were. This led to the question/problem of the nature of memory, which led in turn to the gathering, analysis, and summarization of additional data on amnesias that occurred under a variety of circumstances.

      Ultimately, this research led to the development of a *theory* that there are at least *two forms of long-term memory,* and that they must be different in their anatomical and/or physiological bases. The two forms were labeled *procedural memory* (implicit) versus *declarative memory* (explicit), with the latter having two components, *semantic* (for facts) and *episodic* (for events), and being much more susceptible to disruption. [Squire, L. P. (1987). *Memory and Brain.* New York: Oxford University.]
- Induction can lead to use of the deductive method as a means for further research to enhance knowledge in the subject area.

— Deductive method

- This approach is likely to be used after a *substantial database* has been developed, and thus one or more *general principles or theories* would have already been formulated.
- Deduction moves from the *general to the specific*.
  - ∞ It begins with a generalization — *a general principle or theory* — from which a specific *hypothesis* is logically *deduced* as a potential solution to a *problem* (an answer to a question).

∞ Next it moves through the collection, organization, analysis, and summarization of *observations* (data) in a *test* of the hypothesis and its prediction(s), in order to empirically verify the theory.

- Example of this and another point: There are three major theories for the *acquisition of language,* and they can be evaluated through *strong tests,* i.e., by comparing gathered data against differential predictions of the theories (as noted earlier, this is a very powerful approach to evaluating competing theories).

  Traditional learning theory proposes that the principles of reinforcement and conditioning are the explanations for language development; *cognitive learning theory* proposes that rule learning is involved; and the *biological theory* proposes that there is an evolutionarily built-in capacity from birth for language acquisition and use. These different theories are *testable* by measuring, respectively, the effects of reinforcement and conditioning on language acquisition, the occurrence of rule learning, and the linguistic capacities present at birth. (The previously discussed considerations of *mutual exclusiveness* and *mutual exhaustiveness* should be remembered.)

- Note: Use of the deductive method is likely to lead to the use of a form of *induction* when making judgments about the *implications* of data/results for the theory — thus the two methods are *reciprocal*.

  ∞ When results *confirm the hypothesis* of a study, it is typically concluded that the *general principle or theory* from which the hypothesis was logically derived has been *strengthened*.

  ∞ When the results *do not confirm* the hypothesis of a study, it is typically concluded that the *validity* of the general principle or theory from which the hypothesis was deduced has been *weakened* — note that in either case the conclusions represent *induction* from the specific results to the general theory.

— Strengths and weaknesses of the two methods

- Investigators do not always agree on which approach would be best in a given *research area* at a given *stage of data acquisition*.

  ∞ It might be argued by some, e.g., that the *database* is not yet adequate to warrant the *development of theories*, and that a switch from the *inductive* to the *deductive method* might divert investigators away from the discovery of empirical relationships that are *not implied* by the prevailing theories that had been developed.

  - In other words, *too early a switch* from the inductive method could *bias* the subsequent application of the deductive method.

- Inductive method strengths

- ∞ Allows development of knowledge in subject areas that are *unexplored or poorly explored*, and thus when good general principles or theories have not yet been formulated.
- ∞ Investigators are less apt to be bound by some *conceptual system or theory* when gathering and interpreting data.
  - *Expectations* about the nature of results, in other words, are less likely to occur.
  - *Preparedness* to recognize and accommodate findings as potentially significant is thereby greater, in contrast to the likelihood of rejecting as unimportant certain results just because they don't *conform* to expectations.
  - *Opportunities* for new discoveries are thus maximized.

- Inductive method weakness
  - ∞ Tends to leave knowledge *unsystematized*, at least for a time, while sufficient data are collected, organized, analyzed and summarized for the formulation of one or more general principles or theories.
    - Therefore, the scientific literature provides *less direction* to investigators in generating new research studies.
- Deductive method strength
  - ∞ *Theories* direct future research by suggesting problems or hypotheses that need to be investigated.
- Deductive method weaknesses
  - ∞ Theories can *limit the problems* scientists choose to investigate (as noted earlier).
    - Therefore, potentially *more important problems* might not be investigated.
  - ∞ Theories also create *expectations* about the nature of the results.
    - Hence potentially important findings might not be recognized or accepted (correctly interpreted) simply because they are *not suggested* by the theory and its derived hypotheses.
- Note from the preceding how the *weaknesses* of the deductive method are the *strengths* of the inductive method, and vice versa — in other words, the two approaches are *complementary*.

Review Summary

1. There are seven *primary steps* of the scientific method:
   a. Formulating a problem

   b. Formulating a hypothesis
   c. Designing a study
   d. Collecting and organizing the data
   e. Summarizing and statistically analyzing the data
   f. Evaluating the results and drawing conclusions about the hypothesis
   g. Communicating the findings of the study
2. There are three *secondary steps* of the scientific method:
   a. Generalization
   b. Prediction
   c. Replication
3. There are three ways in which *research problems* become evident:
   a. Contradictory results are found in two or more studies attempting to answer the *same* question.
   b. Gap in knowledge becomes apparent as a result of what has been learned.
   c. Fact needs explaining since it does not fit in with existing knowledge.
4. There are four *basic types of problems* — classes of questions regarding behaviors or mental activity.

   Two of these involve *proximate* causal mechanism: 1) What are the *immediate causes* in terms of external environmental factors and internal psychological, physiological, and anatomical factors, and 2) what are the *developmental factors* in terms of heredity and experience (Nature-Nurture Problem)? The other two types of problems involve *ultimate or distal* causal mechanisms: 1) What is the *adaptive significance* (function) in terms of promoting survival and reproduction, and 2) what was the *evolutionary history* for the species?
5. To be *meaningful* a research problem must be *capable of being answered* with the *tools available to the scientist*.

   Four reasons for research problems being *meaningless or presently meaningless are*: 1) there might be *no empirical reference*, in that it is impossible to obtain relevant objective data since the problem does not deal with observable properties; 2) it might be *presently impossible* to obtain relevant data, although the problem might be *potentially meaningful*; 3) the problem might be *unstructured and vaguely stated*, so that it is not clear what the relevant variables/events are (often correctable by being more specific); or 4) the terms might be *inadequately defined* — a specific source of vagueness (correctable by using operational definitions).
6. *Basic research* investigates problems relating to *fundamental questions* about the principles and mechanisms of phenomena.

   Applied research, in contrast, investigates problems whose solutions are expected to have *immediate application*. Research on both types of problems is essential, however, since basic research establishes the foundation for applied research, which in turn can improve our lives.

7. A *meaningful experimental hypothesis* is an *experimentally testable* statement of a *potential relationship* between independent and dependent variables that *might be the (or a) solution* to a research problem.

   Thus it is more than a *hypothesis in general*, which is simply a *potential answer* to the question put forth by a problem.

8. The *best ways to write a hypothesis* would be as a *conditional statement* ("If ___ then ___") or *quantitatively* as a functional relationship.

9. *The principle of parsimony* is a basic rule and assumption of scientific thinking that states that the hypothesis that is *least wasteful of resources* is to be preferred.

   There are two components of this principle: *Occam's Razor*, that *general explanations* (those that can explain more results) are to be preferred; and *Morgan's Canon*, that the *simplest explanation*, or the one involving the *fewest assumptions*, is to be preferred when it accounts equally well for the same amount of data. Thus, one should seek the *smallest* possible number of *simplest* principles that can successfully explain the *greatest* number of phenomena studied.

10. *Theories* are general propositions or sets of propositions used to *explain* phenomena.

    They differ from hypotheses in that they usually have *more supporting evidence* and are *more general*. The two major *functions* of theories are to *organize established facts*, and thus to arrive at a *systematic body of knowledge* — and hence *description*, and to *guide future research* by generating testable *predictions* — which can lead to *explanation*. As for hypotheses, theories can be scientifically meaningless, or they can be *meaningful,* i.e., scientifically *testable*. Theories are *evaluated* on the basis of how completely and accurately they fulfil their two functions.

11. Different theories/hypotheses about the same phenomena are not necessarily *mutually exclusive*. In other words, they could both be correct — each explaining *different aspects* of the same phenomena, or *different subsets* of the data. A complementary consideration is that a proposed theory (or set of theories) might not be *mutually exhaustive*. That is, there might be *additional or better explanations*.

12. The *inductive method* and the *deductive method* are two *variations* on the general scientific method. They are associated with the two different *functions of theories*: induction is used for *organizing/systematizing* established facts, and deduction is used for *guiding* future research through predictions. They are thus *complementary strategies*.

    The *inductive method* moves from the specific (empirical observations or reflections) to the general (principles or theories) and is more likely to be used when studying topics that have received *little previous attention*, and thus before good general principles or theories have been formulated. The *deductive method* moves from the general (principle or theory) to the specific (empirical observations) and is more likely to be used after a *substantial database* has been developed, and thus when one or more general principles or theories have already been formulated.

13. The inductive and deductive approaches are *strongly interrelated*.

Although they proceed in opposite directions, one often leads to the other. *Induction* can lead to use of the deductive method as a means for further research. On the other hand, *deduction* is likely to lead to the use of induction when making judgments about the implications of results for a theory. Moreover, the *weaknesses* of one are the *strengths* of the other. With the *inductive method*, investigators are less apt to be bound by some *conceptual system or theory* when gathering and interpreting data, but knowledge is left *unsystematized* while sufficient data are collected to formulate general principles/theories. With the *deductive method*, theories are used to *direct* future research, but they can *limit* the problems chosen to be investigated.

Review Questions

1. What are the seven *primary steps* of the scientific method?
2. What are the three *secondary steps* of the scientific method?
3. List and explain the three ways in which *research problems* become evident.
4. List and describe the four *basic types of problems* — classes of questions — regarding behaviors or mental activity.
5. List and explain four reasons for *research problems* being *meaningless or presently meaningless*.
6. Distinguish between *basic/pure research* versus *applied research*, and explain the relationship between them.
7. Define a *meaningful experimental hypothesis* and contrast it to a hypothesis in general (i.e., state the distinguishing characteristics).
8. What are the two best ways to *write hypotheses* (give an example of each)?
9. Describe the *principle of parsimony* and its two components.
10. Describe the two *major functions of a theory*.
11. Explain the difference between *mutually exclusive hypotheses* and *mutually exhaustive hypotheses*.
12. Describe both the *inductive method* and the *deductive method*, and indicate when each is most likely to be used.
13. State one *strength* for the inductive method and one for the deductive method, and discuss the *relationship* between these two approaches.