CHAPTER 3

Scientific Research

Introduction

This chapter shows the intricate nature of research in science. While there are certain, specified steps of a scientific method for conducting an investigation, research is more complex. Actual research involves many factors and revisiting of stages after finding out results of an investigation. As described in another chapter, for example, formulating a medical diagnosis requires changes of strategy and renewing "hunches" about symptoms. In science, there is always a need for reasoning about ideas, both inductively and deductively, to successfully conduct the investigational research.

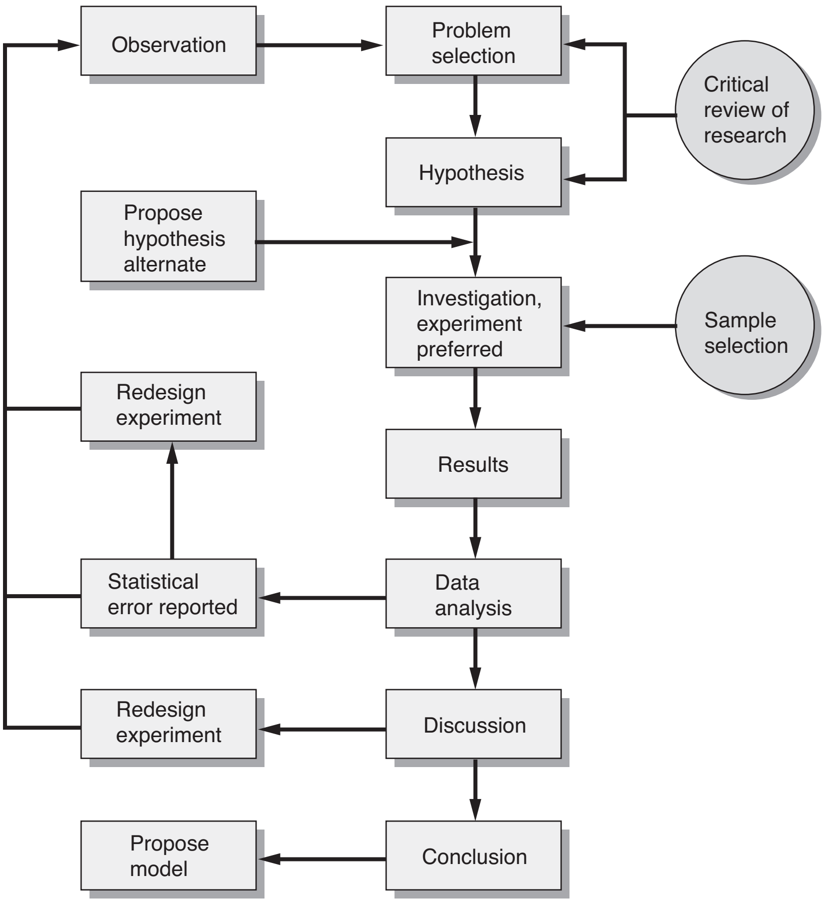

Figure 3.1 Steps of the Scientific Method

The general steps of the *scientific method* are given in **Figure 3.1.** The scientific method is defined as a set of systematic procedures for organized observation and theory building. It is the way in which scientific knowledge is gained.

Observation and Selection of a Problem

Studying the real world and its phenomena naturally leads human beings to ask about some of life's unanswered questions. Children constantly ask about natural phenomena: "Why is the sky blue?" "Why is the moon following us?" A child's curiosity is the starting point for any good scientist. In fact, a child's education is learning about what is known and asking the right research questions about what is not known.

This natural inquisitiveness drives science research through our society and our lives. While we think we know that the moon does not "really" follow us, why and how is the moon moving "with us"? There are many guesses and there is a lot we don't know about the moon. The answers lie in our ability to observe the obvious. The first step in scientific research is to make *observations*.

Then, a *selection* of a research topic to study leads to a background review of the literature about the ***research problem.*** A research problem is the object or process that needs to be studied. In the case of lunar applications, the existing knowledge needs to be explored to determine where society left off regarding moon chemistry and physics. A ***critical review*** of the literature requires that the topic be looked at from every angle. "Critical" means just that—to evaluate inconsistencies in reported research results, find flaws in their designs, identify alternative routes to study topics and critique the mathematical findings. To be a scientist is to be critical of existing knowledge. In this way, new perspectives to research questions are studied and new discoveries are made.

In 1985, a research study overturned the way peptic (stomach) ulcers are studied and treated. A scientist in Australia, Barry Marshall, proposed that the bacterium ***Helicobacter pylori*** caused, and was not a result of, gastric ulcers. Existing research said that these ulcers were brought about by stress and diet and that an ulcer may contain bacteria but only because the ulceration was a place conducive for bacterial growth. In other words, ulcers were not caused by bacteria.

The reluctance of his colleagues to accept the idea that *H. pylori* caused ulcers provoked Marshall to act. In an unprecedented move, the scientist made himself a human guinea pig and ingested a vial of active *Helicobacter pylori.* Marshall recalled in a recent interview:

> Those were frustrating times for me. Most of the experts believed that the presence of *H. pylori* in those who turned up with ulcer problems was just a coincidence. I planned to give myself an ulcer, then treat myself, to prove that *H. pylori* can be a pathogen in normal people. I thought about it for a few weeks, then decided to just do it. Luckily, I only developed a temporary infection.[1]

To the surprise of the scientific community, he formed a number of irritations in his stomach. This showed that his hunch was right—that bacteria did cause gastric ulceration. Whether or not his methods were advised or ethical is another matter.

Hypotheses

The above example shows how after a research problem is identified a ***hypothesis***, or educated guess, is developed and tested. A hypothesis is defined as a possible explanation for a natural phenomenon. In the case of the ulcer, the hypothesis was that *H. pylori* caused gastric ulcerations. It is a conceptualization of the research problem relating the facts to a testable question. The hypothesis should make sense and be based on reasoned research.

There are some general guidelines for evaluating a hypothesis. First, it must be *empirically testable*, meaning the results must be measurable. Second, it must be *logical*. That is, it cannot contradict itself. Third, it must address a question about a *natural phenomenon*. It cannot explore pseudosciences that are immeasurable and within the realm of belief. UFOs and acupuncture can be studied by suitable hypotheses, for example, but must reflect reality by giving measurable results. Fourth, it should seek to explain and *further science*. A hypothesis is an unchecked idea and is only a starting point in the scientific method. Tests of the idea are the true measure of the value of any hypothesis. A hypothesis is only a guess and as such is very subjective.

Scientific Investigations

There are several approaches that can be taken to investigate natural phenomena. Naturalistic observations, scientific modeling, experimental and nonexperimental research will be discussed as possible methodologies. ***Naturalistic observations*** are studies describing how something works in a natural, real setting. Describing a chipmunk's behavior in a forest is a naturalistic observation. It gives the best picture of how the animal works in its normal everyday life. Opposite of this, ***scientific modeling*** predicts results based on simulations of the real world. Models are developed because there is no way to observe it in nature. An example would be a model predicting a hurricane's course. Because the event has not happened yet, it needs to follow predictable patterns based on certain assumptions about the way things usually work.

Nonexperimental research studies general features of events, categories, or places. It is observational and does not seek to intervene in what is being observed. This approach is useful, for example, to map a lake bottom or count the number of fish species within the lake. It is informative and lays the ground work for future studies. These sets of investigation methods will be treated in more depth later in the chapter. However, only the experimental design, described in the following section, which limits subjectivity and bias, is acceptable within the scientific method.

Experimentation

The plan for an ***experiment*** should be detailed and organized and account for as many intervening problems as is possible. What is an experiment? It is an investigation that is organized in a certain way. As discussed in other chapters, it has a *control* and an *experimental* group(s) (that is treated with the studied variable), and *independent* and *dependent variables*. Again, an independent variable is controlled by researchers and a dependent variable is the result of a study. In this way, the influence of the independent variable on the dependent variable can be best determined. Through keeping as many variables as possible constant (not changing) during the experiment, the results are more robust. That is, the dependent variable truly "depends" only on the changes of the independent variable.

Consider an experiment in which reaction time is tested according to age. Our hypothesis states that the time it takes to react to stimuli declines as we age. A simple experiment can be set up in which different age groups are required to "catch" a ruler dropped vertically within the grasp of their hand. The researcher may record how many centimeters of drop it took for each subject to catch the ruler.

What safeguards could be put into place to assure that the experiment is controlled? The same procedure for dropping the ruler, the same conditions of light and temperature for all subjects, and perhaps even the same tone of voice should be used while the researcher drops the ruler. Clearly, there are many variables to control while conducting a sound experiment. Otherwise, ***extraneous variables*** may influence the results of the experiment. Extraneous variables are those that influence the results of an experiment but are not meant to. They are not being studied but may interfere with the influence of the independent variable on the dependent variable. In the reaction time experiment, lighting conditions are one example of an extraneous variable that could affect eyesight and, therefore, reaction time. Careful preparation before an experiment is conducted can minimize the effects of extraneous variables and make the results of the study more valid. This is termed "controlling" the experiment. **Figure 3.2** shows a good experimental design.

Sampling

Usually, it is not possible to study whole populations. In order to study a whole population effectively, a sample or smaller subgroup of the population should be chosen carefully. In an analysis of cholesterol levels in the U.S., an individual cholesterol measurement taken for each member of the nation's population is not economically feasible. Also, studying a single subject's cholesterol test does not provide results applicable to the general population. Therefore, thorough and careful selection of a sample is important. Perhaps taking into account different age groups, fitness levels, and weights would provide a population sample with blood test results more reflective of the overall population.

Many studies do not use the term *subject*, which is usually reserved for humans, but instead refer to a ***sample group*** indicating human and nonhuman units for study. Some sample groups are even geographical regions. For example, when looking at mercury deposition to New Jersey lake sediments, the sample group is a set of

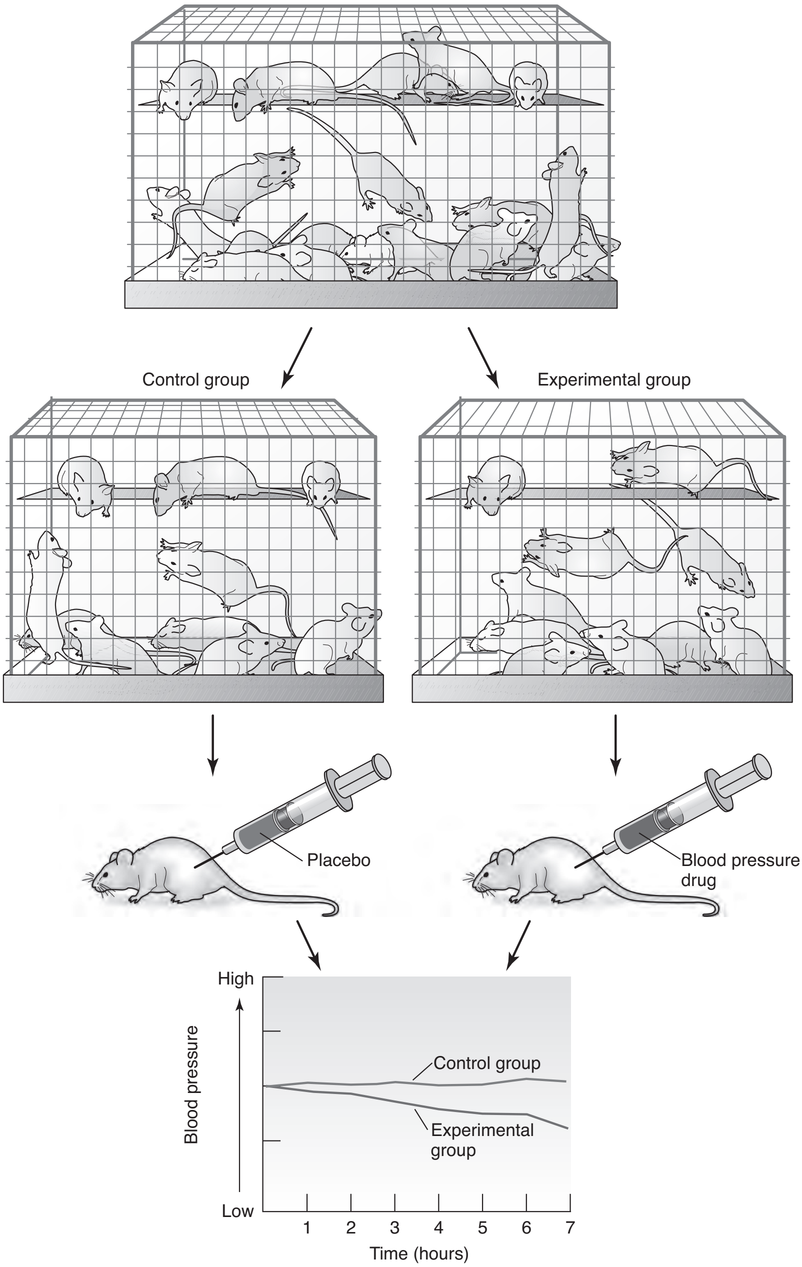

Figure 3.2 Experimental Design: Rats are placed into two groups: control and experimental. Both groups are treated the same except for one variable. In this case, the experimental group is given a blood pressure lowering drug. The control group is given an injection of physiologic saline, the solution the drug is dissolved in (and should have no effect).

measurements from different locations from which soil and sediment samples are collected for chemical analyses. The sample group should always be representative of the larger population.

Bias

The results of any experiment may be flawed due to various types of *bias*. Bias occurs whenever objectivity is compromised. Bias is a mental leaning or inclination, partiality, prejudice, or bent to a researched problem. It gets in the way of obtaining the truth about a phenomenon.

Sampling Bias

Sampling bias occurs when the sample is somehow not representative of the general population. In the above example, bias would occur if samples were collected only from urban areas of New Jersey. The right set of sample areas (e.g., urban, rural, near industrial sources, or in small towns) need to be selected to create an accurate view of mercury deposition representative of the entire state. Sampling only rural regions might result in lower reported rates of deposition, not truly representative of the state as a whole. Thus, careful selection of subjects or sample groups is vital to maintaining the generalizability of any experiment. It is usually most representative to choose random subjects from a whole population. This corrects for possible sampling error or bias.

Psychological Bias

Natural studies present another type of bias problem. As stated earlier, the type of study looking at real life phenomena in nature is termed *naturalistic observation*. It occurs when the investigator gathers data in as natural a setting as is possible. Observations in nature can become unnatural. A chipmunk will not act naturally if it knows it is being observed. In fact, humans often do not act naturally when they are observed. An effect on results due to the study being conducted is termed ***psychological bias***. When limiting this bias, a ***blinded study*** is conducted so that the subject does not know it is being observed. In medicine, often a ***placebo*** (nonpotent) is given in place of a drug, without the subject knowing which it is taking.

Experimenter Bias

Conversely, the experimenter could treat subjects differently based on the groups they are placed in. This form of ***experimenter bias*** can lead to an experimenter unconsciously reporting results more favorable to their particular viewpoints. Keeping secret the type of treatment the subject is getting reduces this form of bias. This is accomplished by a ***double-blinded*** study, in which both the subject and the experimenter do not know in which group the subject is placed.

Data Analysis

Information collected from any study needs to be analyzed systematically and objectively to form conclusions. A variety of research methods may be used to evaluate data. Data analysis may be either quantitative or qualitative in nature.

Quantitative Analysis

Quantitative analysis is defined as the reporting and use of data that are numerical in scope. It is the more traditional scientific analysis system and depends on numbers to find patterns and draw conclusions in the results of a study. The main benefit of quantitative analysis is that it reveals relationships between variables that allow generalization of the results to a larger population. Often, this requires a large number of individuals or units in the sample. To illustrate, consider observing the effects of a drug on rheumatic hand joints. Large numbers of individuals would need to be sampled to allow for adequate statistical analyses. Statistics is the study of the collection, organization, analysis, and interpretation of data. For example, it would help scientists determine if the effects of a drug on improving joint damage is significant. If so, it can be recommended for arthritis patients. Quantitative data analysis is the focus of another chapter. Without math, science has little power to make recommendations or generalizations. Quantitative analysis is what separates science from the many forms of pseudoscience.

Qualitative Analysis

On the other hand, *qualitative data* analyses have fundamental importance in adding richness of detail to scientific studies. ***Qualitative analysis*** is defined as the reporting and use of data that are nonnumerical in scope. It usually studies very few subjects or data pieces but looks at those in greater depth than quantitative research. These studies generally have smaller sample sizes and thus lack generalizability. To illustrate, a single study looking at an individual's rheumatic hand joint and the psychological effects of drug therapy on that person can provide important insights into the drug's overall usefulness. It may look into the patient's personal life, journals and diaries, and effects on the family life to illustrate.

Quantitatively analyzed, an x-ray of the joint may show marked improvement in the rheumatic hand after taking the drug, placing a subject in the "success" group. Qualitatively, however, perhaps the person suffered terrible sleep deprivation and depression resulting in a diminished quality of life. Only through a more in-depth approach would these kinds of results be reported and considered. This is the strength of the qualitative analysis.

Naturalistic observations, or "watching" the subject or setting in its natural proceedings, generally use the qualitative type data analysis. Keeping a journal, taking copious notes, and interviewing are methods of qualitative data collection. For human subjects, conducting focus groups, in which subjects meet in an organized manner to discuss certain topics, allows the researcher to observe in a seminatural setting. In this way, subject interactions and sometimes confrontation may be studied. Data from qualitative studies have a systematical method of coding for detecting patterns and making conclusions.

Results

Basic analysis of the data and reporting of the numbers are placed in a ***results*** section of any investigation. Reporting results is meant to be objective, statistical and numbers-driven. Results are presented in a pithy manner, with little room for elaboration. The chapter on mathematical inquiry will discuss the detailed, quantitative results analysis, which is placed in a results section of a study. Results are never more than straight reporting of the facts. Even in qualitative data reporting, results are trend-based and point to conclusions.

However, there is often pressure on the scientist to present ***positive results.*** That is, results that support the hypothesis. It is often easier to publish a study with a supported hypothesis. Publication of positive results may come with rewards of prestige and money, whereas negative results (refuting original hypotheses) have less impact. Proper reporting of negative results is vital to science, because it ends or limits false leads and drives research into different directions. Falsely representing data and results is highly unethical, as will be discussed in the chapter on scientific integrity. Results should always be objective and unassuming so that readers can draw their own conclusions, possibly conclusions overlooked by the original researcher. In this way, science progresses more successfully with community input.

Discussion

The results are ***objective*** and the hypothesis is ***subjective*** and so the ***discussion*** returns the investigation back to the subjective realm. The discussion interprets the data, explaining it based on the accepted literature as well as the intuitions of the scientist. It is the part of the investigation that is most creative and perhaps even speculative. However, it must be embedded in valid information; the information derived from the results section. The discussion answers questions as to where future research should head, confirms or refutes what is already known, makes alternate explanations for the data, and essentially shows the thinking of the actual scientist.[2] Conclusions are drawn from the analysis of the results in the discussion section.

Scientific Modeling

When actual phenomena are difficult to study, a model is formed about the way something works. As stated earlier in the chapter, a scientific model is a simulation of the real world. A model may derive from an investigation or series of studies. Models make predictions about how something will work by extrapolating from the artificial condition to the real world. Global climate models give predictions about how things will be in the future given current atmospheric conditions. We have no way to see into the future. Thus, the predictions are based on mathematical or artificial situations.

Cancer is an abnormal and uncontrollable production of cells, forming growths called tumors. Tumors may spread to other parts of the body and lead to death. For the past 2,000 years, a group of people in Lin Xian, China, about 250 miles (402 km) south of Beijing, have been dying at high rates (one in four) from cancer of the esophagus, which is the muscular tube moving food from the throat to the stomach. Scientists investigated the Lin Xian area and discovered that the food being grown by the people was low in molybdenum, a soil nutrient for plants needed in small amounts. This caused the crops to concentrate nitrates from the soil to make up for the low molybdenum levels and reduced the vitamin C produced by the plants. Nitrates in the plants were being converted to nitrites in the stomachs of residents; and nitrites are linked to various digestive cancers, including esophageal. It is also known that low levels of vitamin C promote the conversion of nitrates to nitrites. When people were given vitamin C tablets, their production of nitrites dropped. Also, to help reverse the crop production of nitrates, the villagers coat their corn and wheat seeds with molybdenum. As a result, nitrite levels in vegetables have dropped 40% and vitamin C levels have risen 25%. The linked relationships to the villagers' cancer problem are an example of a model, based on scientific investigation, which reveals truth about a phenomenon without actually manipulating variables. It is a model for the workings of the Lin Xian esophageal cancer phenomenon, as shown in **Figure 3.3.** It makes only predictions for how the cancer develops based on the data and model derived. However, it is too early to tell whether or not cancer rates will actually change.[3]

Modeling is an important part of science but is limited due to the removal of real world manipulation of variables. After all, science can cure cancer in mice with anti-angiogenic drugs but in humans the model becomes sketchier. Mice are easier to study because, ethically, more can be done to reduce experimenter errors and eliminate extraneous variables. But in humans, the treatments for cancer don't always work the same. While mice have some biological similarities to humans, they are quite different. We cannot ethically control all aspects of human subjects' lives to eliminate extraneous variables. One group may be eating a different diet or be genetically unlike the other, thus conflating the results of the control vs. experimental group. The models developed from research on mice are sometimes frustrating to the medical community when applied to humans. Despite this, models are important for guiding research. The quote below by Kenneth Paigen expresses this.

> People have a basic misunderstanding of the mouse. They get upset that the exact pathophysiology might be different, but a mouse is not a total mimic. A mouse is a discovery tool, a device, to understand the molecular pathways

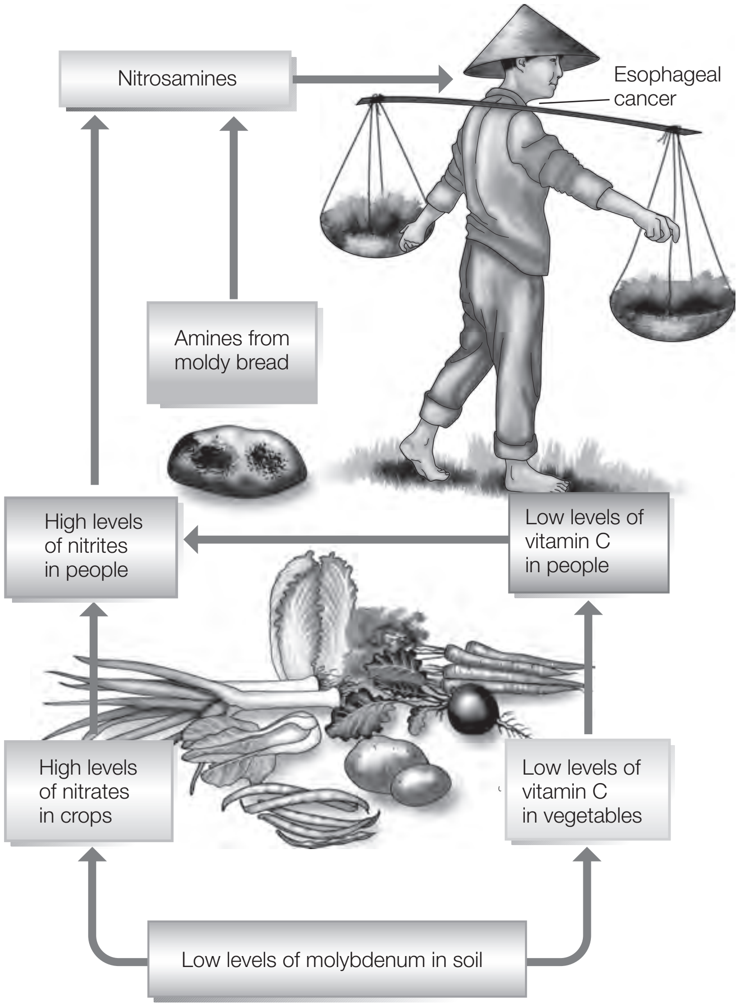

Figure 3.3 Tracking a Killer

> that underlie disease processes. You can't do a lot of types of experiments on humans. You can't order people to mate… We have to turn to models.[4]

Nonexperimental Research

You may have noticed that I have been careful not to use the word "experiment" capriciously in this chapter. A better term is "study" or even "investigation," if the research does not conform to the experimental design. Many studies are *nonexperimental*.

Nonexperimental studies are descriptive: They describe a natural phenomenon. They do not get set up by isolating a variable and having a control group. Nonexperimental studies and investigations are not meant to test predictions, as seen in hypothesis testing and experimentation. Exploring the coast of Iceland for levels of pesticides, describing the Martian soil and topography, and counting species in the Antarctic Peninsula are examples of nonexperimental research. These studies are an inductive gathering of information and they are science. Each furthers our understanding of the universe. Nonexperimental research is foundational in the history of science as will be discussed in other chapters. It is how science began in our history and is still the beginning phases of any scientific investigation, both experimental and nonexperimental.[5]

Analyzing Research: The Incredible Potato

The best way to understand this chapter on scientific research is to read about and analyze an investigation. Consider the study in **Box 3.1** from the *New England Journal of Medicine* reporting a nutritional analysis of different foods in peoples' diets and their resulting weight gain or loss over a period of time. Please pay particular attention to the parts of scientific investigations that give it the strength to make valid conclusions.

In consideration of the last question asked, a host of news agencies reported the results to the public. A *Wall Street Journal* article appeared within a week of the study entitled, "You Say Potato, Scale Says Uh-Oh," which implicated potatoes as a major cause of weight gain and cited in its front cover photo that "boiled, baked or mashed potatoes correlate with a 0.57 pound weight gain" per year for subjects studied.[6] The article implied that even the nonfried potatoes are bad for people. Surprisingly, when investigating further outside of the article facts, the nutritional information on potatoes is very favorable, indicating that they are very healthy to eat. Consider that a boiled potato, cooked in its skin without salt and butter has only 68 calories; is very low in saturated fat, cholesterol, and salt; and is also a good source of vitamin B6, potassium, and Vitamin C. It is the added butter and cream in mashed potatoes or the sour cream and butter on baked and boiled potatoes that likely leads to the weight gain cited.

However, the article omits these extraneous variables co-occurring with the potato-weight gain link. What is the motivation? Is it merely oversight or is it part of a larger groupthink to eliminate carbohydrates from diets? Did this give the newspapers fodder for their anti-carb campaign? Many other news outlets reported the same misleading information about the results of the study, which indicates that it is not a lone error or biased reporter but part of a larger movement. As shown in the example above, the media's reporting of scientific information is often truncated and the public may be easily misled about scientific research results.

BOX 3.1 ARTICLE ANALYSIS: THE INCREDIBLE POTATO

Title: Changes in Diet and Lifestyle and Long-Term Weight Gain in Women and Men, an adapted summary of research by:

Background

While eating less and exercising more is advised, eating different types of foods may contribute to weight gain and weight loss. Specific dietary food choices may be more important than total calories in determining weight and fitness.

Methods

A sample of 120,877 U.S. women and men were studied in terms of their dietary intakes. Samples were separated into three distinct cohorts. Subjects were free from chronic diseases and not obese at the baseline of the study. Follow-up periods were conducted at four-year intervals to evaluate lifestyle factors with weight changes. Adjustments to the data were performed in terms of age, body-mass index, as well as other factors. Various statistical tests were performed to evaluate effects of sex and grouping of the subjects to determine these influences.

Results

The following results were found in each of the four-year period intervals: Subjects gained an average of 3.35 lb, overall (5th to 95th percentile, –4.1 to 12.4); Dietary factors most strongly associated with weight gain were: intake of potato chips (1.69 lb), potatoes (1.28 lb), sugar-sweetened beverages (1.00 lb), unprocessed red meats (0.95 lb), and processed meats (0.93 lb); Dietary factors most strongly associated with weight gain were: intake of vegetables (–0.22 lb), whole grains (–0.37 lb), fruits (–0.49 lb), nuts (–0.57 lb), and yogurt (–0.82 lb); General lifestyle factors were also associated with weight change: increased levels of physical activity (–1.76 lb across quintiles); alcohol use (0.41 lb per drink per day), smoking (new quitters, 5.17 lb; former smokers, 0.14 lb), sleep (more weight gain with <6 or >8 hours of sleep), and television watching (0.31 lb per hour per day).

Conclusions

Specific dietary and lifestyle factors are independently associated with long-term weight gain, with a substantial aggregate effect and implications for strategies to prevent obesity. (Funded by the National Institutes of Health and others.)

In the investigation in Box 3.1:

1. What is the independent variable(s)?
2. What is the dependent variable(s)?
3. Describe the control group(s)?
4. What kind of evidence was collected in the study?
5. What is the hypothesis of the researchers?
6. Is this a quantitative or qualitative study? Why?
7. What is the conclusion of the researchers?
8. Are there any extraneous variables?
9. Do you see flaws in the research? Is there bias in the reporting?

Sources: Dariush Mozaffarian, M.D., Dr.P.H., Tao Hao, M.P.H., Eric B. Rimm, Sc.D., Walter C. Willett, M.D., Dr.P.H., and Frank B. Hu, M.D., Ph.D. *N Engl J Med* 2011; 364:2392–2404, June 23, 2011

KEY TERMS

bias
blinded study
critical review (of research)
discussion
double-blinded study
experiment
experimenter bias
extraneous variable
Helicobacter pylori
hypothesis
naturalistic observation
nonexperimental research
objective
placebo
positive results
psychological bias
qualitative analysis
quantitative analysis
research problem
results
sample group
sampling bias
scientific method
scientific modeling
subjective

PROBLEMS

1. Describe the process by which a research study is selected.
2. "A hypothesis is only an educated guess." Discuss the extent to which this statement is true.
3. Compare the three types of bias. How does awareness reduce bias?
4. What is an extraneous variable? How does it limit statistical power and experimental design?
5. Compare and contrast the following set of terms (give one way the terms are the same and one way the terms are different):
   a. Quantitative and Qualitative Research
   b. Results and Discussion
   c. Nonexperimental and Experimental Research
   d. Control and Experimental Group
6. Which is the most objective in an investigation and why: Results, Hypothesis, or Discussion?
7. What is a scientific model? What are its benefits and limitations in scientific progress?
8. How is nonexperimental research important to scientific progress?

REFERENCES

1. National Institutes of Health: Office of Science Education. Available online at http://science.education.nih.gov/home2.nsf/Educational+ResourcesResource+FormatsOnline+Resources+High+School/928BAB9A176A71B585256CCD00634489. Accessed July 6, 2012.
2. Keppel, G., Saufley, W., and Tokunaga, H. 1992. *Introduction to design and analysis: A student's handbook*, 2nd ed. New York: W.H. Freeman and Company.
3. Chiras, D. 2006. *Environmental Science*, 9th edition, Burlington, MA: Jones and Bartlett Learning.
4. Lewis, R. 1998. How well do mice model humans? *The Scientist* 12 (21):10–11.
5. Lee, J. 2000. *The scientific endeavor: A primer on scientific principles and practice*, San Francisco: Addison Wesley Longman, Inc.
6. Hobson, K. 2011. *Wall Street Journal*, You Say Potato, Scale says Uh-Oh, June 23, 2011.