Critical Reading 1

C ritiquing research is not something that comes naturally. People (including psychologists) spend many formative years learning and using other ways of appraising truths. They are governed by other ways of evaluating the validity of people's claims before they learn about the scientific method and before they gain an understanding of the methods that are used in psychological research. First learned, deeply ingrained modes of thought continue to affect people as they go about the task of evaluating what they are told orally and in print and to influence them in deciding what evidence to consider valid.

This chapter reviews the various methods that people use for seeking the truth and their differing views on what constitutes proof. These views naturally affect the reader's attitudes toward the material under review and determine whether the material is accepted uncritically or is subjected to critical appraisal. The crux of the critical appraisal, in turn, depends on the rules of evidence that the reader holds necessary and the reader's knowledge of research design.

Methods of Seeking Truth— Proof and the Rules of Evidence

Diverse methods of seeking truth and varied codes of rules of evidence may conflict, coexist, or even work in harmony

with each other for the same individual in different spheres of life. What works for a person in one aspect of life may not be satisfactory in another. Some people cannot "change hats" easily; others can switch seamlessly from one mode of thought to another, as occurs, for example, when one reads a historical novel, interacts socially, goes to church, and reads scientific studies.

FAITH

Authoritative Pronouncements

Some people take the judgment of one or more authorities as sufficient evidence. This can take the form of ex cathedra pronouncements by religious figures or authoritative pronouncements by experts. Experts can be highly reputable scientists whose credentials give cachet and weight to their pronouncements. What if two or more experts disagree, as they so often do? With the unerring benefit of hindsight, one can see how mistaken highly qualified scientific experts have been in their pronouncements. In *The Experts Speak: The Definitive Compendium of Authoritative Misinformation*, Cerf and Navasky (1984) provided retrospectively amusing quotes from outstanding authorities in various fields. The physicist Lord Kelvin, who was president of the British Royal Society, announced in 1895 that "heavier-than-air flying machines are impossible" (p. 236). In 1897 he stated that "radio has no future" (p. 206) and in 1900 claimed that "X-rays are a hoax" (p. 302). In 1878 Professor Erasmus Wilson of Oxford said, "[W]hen the Paris Exhibition closes, electric light will close with it and no more will be heard of it" (p. 203). Einstein is reported as having said in 1932 that "There is not the slightest indication that [nuclear] energy will ever be obtainable. That would mean that the atom would have to be shattered at will" (p. 215). In 1956, the leading British astronomer Richard van der Riet Wooley allegedly said, "Space travel is utter bilge" (p. 258).

Democratic Judgments

Placement of faith in the judgment of many in contrast to the dicta of single authorities is a common way of deciding what is right. The saying goes, "Fifty million Frenchmen can't be wrong." Confidence is instilled when a judgment is shared by many people. Committees decide what is true (and the blame can be shared among them when they are proven wrong).

REASON

Another way of arriving at truth is by reason alone. Under certain circumstances pure reason can lead to what is called *necessary truth*. In a

priori reasoning, conclusions are deduced from known and self-evident principles or definitions, such as "all men are mortal." The syllogisms of formal logic and the proofs of pure mathematics lead to necessary truth.

However, not all questions can be answered by pure reason alone. If they could, there would be no need to design or carry out experiments. Empirical work calls for an inductive a posteriori process whereby final judgments are based on experience, and principles are generalized from facts that are established by experience rather than by the application of reason alone.

FEELING

Ad hominem arguments appeal to the feelings and passions rather than to reason and intellect. The person asserts subjective truth in saying, "I feel it to be true in my heart, in my guts. Therefore, it must be true." The determination of objective truth from gut feelings is problematic when two people have equally intense but opposite feelings.

SENSORY INFORMATION AND EXPERIENCES

One of the most compelling and unshakable proofs for people is personal experience. There is a certainty to the subjective truth of the statement, "I have seen it, I have experienced it. Therefore, I know that it is true." The reality of extraterrestrial presences is firmly established for people who believe that they have witnessed, sensed, or experienced them. One step removed from such proof is the acceptance of the testimony of others who claim to have sensed or experienced something. Scientists, who are trained to be skeptical, are not immune to the importunity of personal experience. The famous self-deception of Professor René Blondlot (1905), the noted French physicist, who thought he saw spectroscopic evidence of nonexistent "N-rays" being emitted from metals and who convinced the scientific community of their existence, is an instance of this. The fallibility of the human senses is well-known to psychologists who have studied illusions and to magicians who exploit this fallibility to their advantage.

LEGAL METHODS

Legal methods for arriving at truth call on all of the above methods including authoritative testimony of experts, appeals to pure reason, ad hominem arguments, and testimony of witnesses about what they have seen, heard, or experienced. Faith is then placed in the judgment of the jury members who weigh the evidence and decide the truth. A

simple democratic majority is not sufficient for conviction in criminal cases that demand unanimous agreement.

EMPIRICAL AND EXPERIMENTAL METHODS

Empirical (which include experimental) judgments are based on objective and systematic observations or experiments. In experiments, the scientist exercises control over the variables by careful prearrangement of conditions. In nonexperimental empirical work, the scientist observes and records but does not manipulate the variables. These approaches to discovering or establishing truth reflect a basic skepticism about proof and are characterized by a "show me," "prove it," and "what is the evidence?" attitude about alleged "facts." Although the conclusions that one reaches are neither subjectively nor objectively certain, one can judge how probable they are. Even though absolute certainty may be beyond reach, one does aspire to attach a level of probabilistic confidence to judgments about what is true.

Approach to Reading Critically

REVERENCE FOR THE PRINTED WORD

In the days when few people could read, written symbols or words were painstakingly carved on stones, walls, pillars, and clay tablets. Anything that had to go through such a laborious process and was obviously being recorded for eternity must have seemed to viewers to be of utmost importance. Publication in this mode did not invite critical scrutiny, nor did it serve as a prelude to dialogue.

Publication requires either a one-of-a-kind public display or the dissemination of multiple copies. There was an advantage to the use of soft and flexible materials, which could be posted, copied, transported over distances, or stored in libraries. The printing press made it far easier to make alterations. Words could be added or deleted and sentences or paragraphs moved about. Paper itself became less durable, as if to acknowledge that the written word is not permanent, that ideas do change, and that new truths come forward and old ones have to be revised or replaced. Obsolete documents, as distinct from stone, can easily be crumbled or torn up and discarded.

As literacy became more common, and the means of disseminating thoughts through the written word became available to more people, previously entrenched truths were increasingly challenged. People not only could read, but they could privately disagree with, and even

publicly rebut, what they read if the social climate of the time was sufficiently enlightened to permit it. New writing instruments such as lead pencils and quill pens made it possible for ordinary people to write down their thoughts. The eraser, conveniently placed at the end of the pencil, accommodated an expectation that errors of thought and inadequacies of expression would occur. Things written in pencil have a tentative and impermanent air about them. The fountain pen, an invention of the 19th century (Ansley, 1935), liberated people to leave their more indelible mark anywhere. During its brief history, the typewriter made it possible for anyone to produce official-looking printed words and to make rapid changes if desired. Now that computers have become the writer's choice, whole sections can be deleted with the press of a button or moved about thanks to the miracle of the cut and paste function keys.

From this brief and sketchy sweep through the history of writing, a few things stand out. Worthwhile thoughts have not gotten any easier to come by, but it is certainly simpler to record them and to communicate them to others. When almost anyone can print out words and thoughts, there is, I suspect, a decline in reverence for the printed word. This sets the stage for critical reading and makes it a more natural and acceptable thing to do.

CRITICAL READING

The critique of what is presented as a scientific study can justifiably address the way the study does or does not meet the scientific standards for evidence and proof. In reading such a study, the reader's focus is on its scientific soundness, not on whether the findings conflict with preexisting faith, beliefs, or ideas about social acceptability; not on whether the results conflict with expert opinion or clash with other methods of establishing truth and gathering evidence; and not on whether the results are in harmony with his or her gut feelings. The critique mirrors the reading itself.

Even though it cannot provide more than probable truth and must be taken with caution, empirical (and particularly experimental) research is still the best way to look for answers to certain types of questions. Presentation of all of the ways an experiment can founder and the emphasis that this book places on scrutinizing and critiquing research is in no way intended as an indictment of the research process or as an expression of reservation about its value as a method of seeking the truth.

Doing research and reading research reports critically are both arduous processes. It is much easier to say, "This is too hard, let's just brainstorm it and take a vote," or "Let's just ask Dr. Jones, he ought

to know," or "Let's just review the literature and take what the experts say," or "Let's hold a symposium, invite some experts to debate the question, and give the nod to the one who is most persuasive." People may become so overwhelmed with all of the difficulties associated with doing and critiquing research that they detach themselves from it and declare that they are simply not interested in this method of finding things out or that they find it too problematic and too inconclusive. Others admit that it is too difficult and too technical for them to judge by themselves and that they prefer to leave it up to the researcher, who, after all, must know best. They then take, on faith, whatever is printed in the scientific journal. This defeats the whole purpose of presenting research for public scrutiny.

The phrase "research shows" is as revered in some circles as ex cathedra pronouncements are in others. The important thing is for the person who utters these words to have read the research, to have understood it, and to have evaluated it before using its findings to support a position or to inform a practice. When the research is truly sound, however, there is no more powerful way to make a point than to cite research findings.

Critical reading requires a mental set of a particular kind. This mental set can be taught, encouraged, and nurtured. Conversely, it can be discouraged or even forbidden. What is involved, first of all, is a kind of general open skepticism that enables one to bring a "show me" or "convince me" attitude to the reading table regardless of how authoritative the author may be or how attractively the words are packaged. The reader then engages in an interactive dialogue with the manuscript. The critical reader applies exceptional focus to each sentence that is written, contemplates the meaning behind it, and thinks about the thoughts that were not written and wonders why they are absent. This cannot be done in haste, and it often requires multiple rereading and rumination about particularly troublesome sections. Critical reading of scientific materials is very different from recreational reading of science fiction or of a novel (although these too can be read critically by literary scholars). When reading scientific documents, every thought that is expressed, every conclusion that is reached, has to be screened through one's internal skeptical scanner. If everything meets the most stringent cognitive challenges, one comes away enriched and gratified. On the other hand, one may accept some of it but have reservations about other parts. A third scenario is that one may reject it in its entirety and resent having spent the time on it. None of these scenarios is the same as reading in a passive–receptive mode, with one's critical antennae down, and accepting the author's conclusions wholesale.

Uncritical acceptance of conclusions leads to the incorporation of misinformation into one's body of knowledge. It is no trivial matter when such material misinforms public policy or becomes a false guide to one's clinical practice. "Research shows" is one of the favorite expressions of psychologists who are called on by the media to express their professional opinions on a wide range of topics, who are asked to consult with or testify before lawmakers about social issues that affect public welfare, or who are relied on to give expert counsel to other health service providers or to educators. Research psychologists thus carry a heavy burden of responsibility for assuring the accuracy of their claims about their results. In turn, psychologists who cite or apply the research findings of others share that responsibility. They have an obligation to use their critical reading and evaluation skills in reviewing a study before they cite it as evidence that supports a point of view and before they apply the findings in their clinical work. Uncritical acceptance of invalid research can impede the development of the field and jeopardize human welfare.

How different this process of critical reading is from gaining scientific information from summaries, abstracts, or digests, as busy professionals are often tempted to do! These provide conclusions to be accepted on faith alone. Yet nobody has the time to read all of the scientific and professional journals in the field. Digests and abstracts do serve a useful purpose. The best way for readers to use them is as a screening device for subject matter that may be of special interest, to be followed by critical reading of the actual articles. When faced with an original full-length article, experienced critical readers often skim through the article quickly so that they can get a general overview of the contents and can decide whether they are sufficiently interested to read it more carefully. If they decide to engage, they then begin to reread the article in a fully focused critique mode. Sometimes several readings of crucial sections are required. This is particularly true when the text is fuzzy or lacks readily accessible detail.

Interactive Reading

The passive reader is a recipient who takes in what the text says. At the end, the reader may have a general impression of the research and some points of agreement or contention. The interactive reader anticipates what is to come and then discovers whether these expectancies are met along the way. Adoption of the interactive mode places the reader in the best position to evaluate the work critically. Throughout this book, I call attention to the kinds of expectancies that flow naturally from given antecedents:

- that the research question guides the literature review
- that the literature review and the statement of the problem inform the hypotheses
- that the hypotheses set up design expectancies and suggest what variables should be controlled
- that understanding what should be controlled influences the design and the procedures and prompts a reader to look at the methods that are used to exercise control
- that the hypotheses and design and type of data dictate the method of data analysis
- that the analysis of the data influences the kinds of conclusions, inferences, and generalizations that can be made.

The interactive reader is working all of the way, checking evolving expectancies against accomplishments in the study. The critique, whether negative or full of praise, can be focused and justified.

When in the interactive reading mode, the reader selects an article to read because the title sounds as though it might be of interest. Seeing the title, "Effect of Teaching Method on Reading Ability," one expects to find a cause–effect experiment in which the method of teaching is systematically varied, and reading ability is compared in groups of children taught by different methods. The review of the literature and the statement of the problem reinforce that expectation, and the hypotheses are worded in a causal manner. Turning to the Method section, the reader is soon disappointed to find that instead of manipulating teaching method systematically, the researcher has located different elementary schools within a large school system. The method of teaching reading varies from school to school. One immediately wonders about the equivalence of the various groups of students to start with and the equivalence of the skills of the teachers. This sets up yet another unfulfilled expectation that the researcher has made sure that the groups are equivalent and that the teachers are comparable. Proceeding, the reader anticipates a valid and reliable measure of the dependent variable, reading ability. Instead, supervisory rankings are made of the various classes, and testimonials are given about the accuracy of this ranking procedure. Turning to the analysis of the data, the reader expects, in view of the ordinal data, that nonparametric statistics are used. Again, expectancies are not met; the author uses parametric statistics instead. Curious about how all these issues are handled in the Discussion section, the reader finds unjustified authoritative pronouncements and unfounded causal statements in support of the hypotheses.

How different is this interactive reading process from the passive–receptive mode where the reader sees the article as excessively

detailed and would much prefer to read the abstract and skip the rest. Good "critical" interactive reading does not require the uncovering of crucial flaws. A special pleasure and admiration come to the interactive reader when serial expectations are met, and in fact exceeded, when the researcher anticipates and forestalls problems that the reader has not thought about, when the research procedures not only fit basic requirements but do so in a particularly clear and original way, where the data are analyzed in an especially convincing manner, where the inferences from the data are unquestionably sound, and where the discussion is impressively profound and thought provoking.

At one extreme, the critical reader of sound scientific material will say, "I've learned something and will incorporate it into my knowledge base, I'll accept the conclusions and will cite them and apply them when appropriate." At the other extreme, the reader of a flawed study cannot accept any of it as true.

Knowledge of Research Design

Readers may sit down in a state of heightened alertness, eager to critique an article of interest. Are they prepared? Do they know what to look for and on what to focus their attention? Will they see the defects or be oblivious to them? If readers do not have the background and training to critique the article intelligently, they are bound to have misgivings. They are in a position analogous to that of a person who is contemplating the purchase of a used car. Should faith be placed in the salesperson's appraisal of the car's merits? Does the buyer know how to assess it? Does he or she know all of the potential problems that may be present in the car's mechanical and electrical systems and how to evaluate their presence or absence? People who lack the specialized knowledge can always hire an expert to advise them. Many people rely on drama and motion picture reviewers to screen the offerings and to offer suggestions about what is worthwhile. When it comes to scientific journal articles, however, there are usually no reviews to consult. Buoyed by the knowledge that articles in some journals have been previewed by referees and read by an editor, readers may feel confident that the articles have been pasteurized and are safe for consumption. Nonrefereed journals do not even offer the added protection of a preview.

This kind of prepublication review in select journals is a helpful form of quality control. In 1994, the rejection rate of articles submitted for publication in American Psychological Association (1995) journals averaged 75% and ranged from 44% to 92%. Regardless of how stringent this appears, no article that succeeds in being published is accompanied by a guarantee of excellence. How many times has one

seen flawed articles in refereed journals? How many times has one read critiques of an article, followed by the author's rejoinder? Is it not commonplace, and in fact expected, that authors of journal articles critique the work of their predecessors? This would never be necessary if all articles in refereed journals were beyond criticism. Obviously this is not the case.

Any research article has three components: (a) the technical content as reflected in the literature review, statement of the problem, and formulation of hypotheses; (b) the research itself; and (c) the interpretation and discussion of the results. A reader who knows much about a topic, but little about research design principles, is not in a good position to evaluate a research article critically. A reader who knows little about the topic can nonetheless evaluate the design and execution of the research. Such a reader may miss out on some potentially crucial details regarding the choice and use of specialized apparatus or the selection of the best available criterion measures or most advanced techniques. Generally, however, readers who review material that is within their broad general sphere of knowledge should be able to critique any but the most highly specialized articles.

Principles of research design transcend content areas. The effectiveness of a reader is therefore dependent on knowledge of research design and on skillful application of that knowledge. The remaining chapters in part I discuss those aspects of research design that are essential to know about if one is to conduct a comprehensive appraisal of a research article. A systematic approach to the critique of research reports is outlined in the Prologue to part II (pp. 161–167).