The headline of the Jan 16, 2009, article in *Time* magazine by John Cloud exclaimed, “How to lift your mood? Try smiling.” You might respond, Wow! Why didn’t I think of that? Or you might dismiss at first blush (no pun intended) as preposterous the claim made in the headline that smiling can actually lead to improved mood. You might think to yourself, everyone knows our emotions don’t
work this way—we smile as a result of being happy, smiling doesn’t cause us to be happy. However, as we will learn in this chapter, the scientific method can be applied to this chicken-and-egg question as to what comes first, the simple act of smiling or the positive emotional feeling. And what you will learn in this chapter might surprise you, and the claim in this headline could seem less preposterous than you might have originally thought.

As you recall from Chapter 1, there are several techniques available in the research toolbox from which you can choose. The research methods and design you choose depend greatly on the questions asked and the answers you hope to discover. There are two very broad research approaches used by researchers: experimental and non-experimental designs. For example, studies that use naturalistic observation methodology are nonexperimental in that actions and events are carefully measured and catalogued but independent variables cannot be directly manipulated and extraneous factors that can confound results are difficult to control. Nonexperimental designs encompass several types of studies, including studies referred to as quasiexperimental, correlational, survey, and single-subject or small-N research. Thus, if you wanted to examine the relationship of smiling and positive mood using a nonexperimental design, you might have participants watch a comedy film, record the number of times each participant smiled, and then have participants complete a questionnaire on their mood. Suppose you found a relationship between increases in smiles and positive mood. Could you conclude that there is a causal relationship between smiling and positive mood? As you know, the answer to this important question is “no.”

On the other hand, an experimental study can be used when the researcher wants to test cause and effect, such as investigating whether smiling causes positive changes in emotional feelings. In an experimental design, a researcher is able to exercise control over the variable or variables that are assumed to be the causal agent(s) producing the predicted effect. In this chapter, we will learn about the nature of these variables, and we will also learn that causality in human psychological research even with an experimental design is surprisingly complex. But for now, the important message is that only an experimental design allows for testing of cause-and-effect relationships.

THE GOALS OF SCIENCE

Paul Ekman is a world-renowned psychologist who has devoted his career to studying facial expressions of emotions. In the 1960s as a young psychology professor at the University of California at San Francisco (UCSF), Ekman found that most social scientists understood emotional expression in the face to be culturally determined by a set of learned social conventions. He had a very different idea, one that could be traced to Charles Darwin’s 1872 book *The Expression of the Emotions in Man and Animals*, in which Darwin argues that all mammals communicate their emotions through facial expression. Malcolm Gladwell, in a delightful 2002 *New Yorker* piece, described how an undaunted and inspired Ekman set out to discover the rules governing how we express emotions in our faces. Ekman began, Gladwell writes, by challenging the idea that such rules are learned social or cultural conventions.
Ekman first traveled to places like Japan, Brazil, and Argentina, carrying photographs of men and women posing various distinctive faces. Everywhere he went in the developed world, people agreed on what those expressions meant. But because these people all lived in developed countries, they, Ekman knew, could still have picked up the same cultural rules watching the same movies and televisions shows. The indefatigable Ekman, armed with his trusty photographs, next traveled to the most remote villages in the jungles of Papua New Guinea, and there he found, lo and behold, that these tribe folk also had no problem interpreting various facial expressions of emotions. In this very important sense, then, as you recall from Chapter 1, Ekman began his work as a true cultural psychologist.

Over the next half century, Paul Ekman would build a program of psychological research around his breakthrough discovery that facial expressions of emotions were not socially learned but rather are the universal products of evolution. His nonexperimental studies of facial expression across cultures would be extended to elegant experimental studies by him and other researchers (see also, e.g., Leppanen, & Nelson, 2009; Mandal & Ambady, 2004). Together, this research established the human face as an exquisite and efficient organ of emotional communication.

Exhibit 2.1 Universal Expressions of Seven Major Categories of Facial Emotions

Eve Ekman, the daughter of psychologist Paul Ekman, demonstrates the seven major categories of facial expressions. Top row, from left to right: Fear, anger, happiness, contempt. Bottom row: surprise, disgust, and sadness.

Source: Paul Ekman, PhD/Paul Ekman Group, LLC.
Description

The principal objective of research is to provide a scientific understanding of the topic of investigation. Scientific understanding entails two distinct but related processes: description and explanation. We first describe the phenomenon that we intend to study. This may seem obvious, but careful and precise definitions of key concepts and the measures we develop for those concepts are critical for scientific understanding.

Generally, researchers describe and define terms both conceptually and operationally. A conceptual definition provides the meaning, often rather broad in scope, of an abstract term, such as intelligence, anxiety, or emotion. Very similar to what you would find in a dictionary, a conceptual definition demarcates a semantic or linguistic meaning of a psychological term, that is, its usage in words, texts, and language. For example, intelligence as a concept may be defined as the general ability that enables an individual to comprehend the world and to deal effectively with its challenges (Wechsler, 1997). For Ekman, the term emotions represents another key concept that he defines broadly as a rapid and coordinated response system that evolved to enable humans to react quickly and effectively to events that affect their welfare (Ekman, 2003). In this conceptual definition of emotions as a response system, facial expressions are a critical medium of communication.

Operational definitions follow from conceptual definitions. An operational definition indicates how a concept is coded, measured, or quantified. It may be as simple

<table>
<thead>
<tr>
<th>Table 2.1</th>
<th>Operational Definitions for Producing Seven Major Categories of Facial Expressions of Emotions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SADNESS</strong>: The eyelids droop as the inner corners of the brows rise and, in extreme sadness, draw together. The corners of the lips pull down, and the lower lip may push up in a pout.</td>
<td><strong>Seven Major Categories of Facial Expressions of Emotions</strong>&lt;br&gt;<em>(adapted from Duenwald, 2005)</em></td>
</tr>
<tr>
<td><strong>SURPRISE</strong>: The upper eyelids and brows rise, and the jaw drops open.</td>
<td></td>
</tr>
<tr>
<td><strong>ANGER</strong>: Both the lower and upper eyelids tighten as the brows lower and draw together. Intense anger raises the upper eyelids as well. The jaw thrusts forward, the lips press together, and the lower lip may push up a little.</td>
<td></td>
</tr>
<tr>
<td><strong>CONTEMPT</strong>: This is the only expression that appears on just one side of the face: One half of the upper lip tightens upward.</td>
<td></td>
</tr>
<tr>
<td><strong>DISGUST</strong>: The nose wrinkles and the upper lip rises while the lower lip protrudes.</td>
<td></td>
</tr>
<tr>
<td><strong>FEAR</strong>: The eyes widen and the upper lids rise, as in surprise, but the brows draw together. The lips stretch horizontally.</td>
<td></td>
</tr>
<tr>
<td><strong>HAPPINESS</strong>: The corners of the mouth lift in a smile. As the eyelids tighten, the cheeks rise and the outside corners of the brows pull down.</td>
<td></td>
</tr>
</tbody>
</table>

*Source: From “The physiology of . . . facial expressions: A self-conscious look of fear, anger, or happiness can reveal more than a lie detector” by Mary Duenwald. January 2005, Discover magazine.*
as an operational definition of gender in which female is coded as 1 and male as 2 (or vice versa). No single operational definition can capture fully the concept it is intended to measure. An operational definition is among several possible objective and measurable indicators of a concept. We will learn more about the relationships of variables and concepts in Chapter 4 when we discuss psychological constructs.

In the Ekman work, the Facial Action Coding System provides the operational definition of various facial expressions of emotions. The coding system defines specific combinations of facial muscle movements that are universally and discretely generated when certain emotions are elicited (see Table 2.1 and Exhibit 2.2). In essence, specific sequences of facial muscle contractions are predictive of particular emotional expressions. Thus, we see that an important aspect of a scientific description and explanation is establishing *predictive relationships* between two events or occurrences.

**Exhibit 2.2** The Facial Affect Coding System breaks down the muscles associated with facial expressions of various emotions. For example, the orbicularis oculi and zygomaticus are activated to produce a smile, the corrugator is activated during frowning in anger, and the levator is used to produce the grimace of disgust (from Niedenthal, 2007).

Note also that the coding system is transparent so that a researcher can learn these patterns of facial muscle movements and design studies to replicate as well to extend findings about human emotional expression across different cultures. In 1967, Ekman traveled to the highlands of Papua New Guinea to conduct research with the Fore people, who had never been exposed to movies, television, magazines, or many foreigners, but who nevertheless recognized facial expressions of emotions exactly as Westerners did. Through the use of the Facial Action Coding System, Ekman identified seven discrete emotional expressions that humans in different cultures develop a similar capacity to recognize: happiness, sadness, fear, anger, surprise, disgust, and contempt. The generalizability of such findings was especially impressive. You can try your luck at identifying facial expressions of these seven emotions as portrayed by the actor in Exhibit 2.1. You will find descriptions of the particular combinations of facial action units for each expression in Table 2.1.

**Explanation**

Description and explanation are the goals of science. Explanation can include both prediction as well as establishing cause and effect. Causality, however, is not a simple matter. In fact, philosophers since the time of Aristotle have debated the true essence of causality. For our purposes, let’s think of causality as requiring three kinds of evidence, as described by Cook and Campbell (1979). The first kind of evidence is that of *temporal precedence*. This evidence establishes that the cause precedes the effect. Thus, one step toward demonstrating smoking causes lung cancer is data demonstrating temporal precedence: Smoking cigarettes occurred first, and lung cancer, tragically, followed. The second type of evidence needed to establish causality is *covariation of the cause and effect*. This means that when the cause is present, the effect occurs, and when the cause is absent, the effect does not occur. Thus, we need to know that people who smoke cigarettes contract lung cancer, and those who do not smoke do not contract lung cancer. Third, causality requires the elimination of *alternative explanations*. In other words, a researcher must show that nothing other than a causal variable could be responsible for the observed effect; that is, there is no other plausible explanation for the relationship. In our smoking example, suppose a third variable, such as social class, could explain the relationship of smoking and lung cancer. This alternative explanation could come into play if cancer victims were found to come from socio-economic strata different from those of healthy individuals.

As you can see, these three prongs of causality are difficult standards to meet in human psychological research. This holds even with a true experimental design, which ostensibly allows the researcher to test whether changes in an independent variable causes changes in the dependent variable. The problem is that it is impossible in human studies to control for all extraneous, *confounding* (see Chapter 1), or so-called *third variables* that could account for the observed effect.

However, even when we can’t identify causality, a scientific explanation can provide highly useful information that helps to specify the rules under which a certain phenomenon occurs. As an example of a scientific explanation governing facial expression of emotions, consider the Ekman Facial Action Coding System.
In this coding scheme, each of the seven universal facial expressions presented in Exhibit 2.1 can be decomposed into a specific sequence of microaction units as outlined in Table 2.1. For example, the facial expression of disgust can be broken down into microaction units of wrinkling the nose and raising the upper lip while protruding the lower lip (try it yourself).

**Practical Knowledge**

Scientific understanding through the process of description and explanation often leads to interesting practical application. Studies are often categorized as either *basic research* or *applied research*. Basic research addresses fundamental questions about the nature of abstract psychological processes and ideas, such as emotion, intelligence, reasoning, and social behavior. Applied research addresses important questions that are thought to be of immediate relevance in solving practical problems. What television advertisements are most effective in reducing illicit drug use in children? What is the most effective teaching method for learning math? Do early-intervention programs, such as Head Start, lead to better outcomes? Such questions are investigated and studied in applied research. In fact, a major area of interest and study in applied research is called *program evaluation*. Program evaluation studies the effects on behavior of large-scale policy changes and as well social reforms and innovations occurring in government, schools, courts, prisons, businesses, health care, housing, and so forth.

In reality, these distinctions often blur and the basic and applied research designation is probably most accurately viewed as falling along a continuum. Both are essential and of equal importance, with scientific progress depending on the combined efforts and fruits of basic and applied research. Moreover, basic research often has exceptionally important applications that are impossible to predict. For example, the principles and findings from Ekman’s basic research on facial expression of emotion are now applied in a variety of settings. The most high-profile of these settings is public safety, including the CIA, the FBI, and the Department of Defense. But other areas of clinical application are also prominent, as in modulating facial expression of emotions as an important psychotherapy technique in treatment of debilitating mental problems (Linehan, Heard, & Armstrong, 1993). In addition, the June 2008 issue of the *Journal of the American Academy of Dermatology* features an article by Alam and colleagues reviewing evidence that botulinum toxin injections (“botox”) may have an antidepressant effect by reducing frowning and other negative facial expressions. “Can looking better make you feel happier?” is the engaging question that is included in the title and is the focus of the Alam and colleagues review.

**Sources of Research Ideas**

Generating ideas may very well be one of the most exciting aspects of the research process. In this the first step of the research process, our mindset may become more expansive, curious, and creative, as we would like to generate novel ideas or new ways to think about old problems. While the scientific method provides us
with a prescriptive, disciplined, and objective format to test the veracity of our ideas, the question for us to consider now is how we actually come up with the ideas in the first place.

### Starting With Observation

Suppose you have no formal experience in research but you have always been curious about the human face. In many ways, you have chosen a wonderful topic for research that can be studied by the power of simple observation. Science often begins with simple observation, which can serve as a source of both evidence and ideas. Charles Darwin, for example, generated the theory of evolution by natural selection exclusively on the basis of simple observation. In his later 1872 book *The Expression of the Emotions in Man and Animals*, and again relying exclusively on observation, Darwin would make the case that all mammals regularly display emotion in their faces. For us mere mortals, we might look for the “agony of defeat and the thrill of victory” etched on the faces of the athletes when watching our favorite sport. See for example Exhibit 2.4 from Sandy Nicholson’s (2008) book *2nd: The Face of Defeat*, which in many ways captures the power of observing facial expression of emotions.

Often in research, our observations will need to be broken down into simpler units that allow for more precise quantification and measurement. Indeed, as we have learned about the research of Paul Ekman, the complexity of human emotion is much easier to study by making it simpler. At the risk of hyperbole, how elegantly simple and brilliant is the Ekman system of decoding various facial expressions into distinct sets of specific muscle movements? In a similar vein, as we will see later in this chapter, Strack and colleagues showed how two simple positions of holding a pen either gently between the teeth or tightly between the lips induced facial expressions of smiling and frowning, respectively (Strack, Martin, & Stepper, 1988).

Observations often act to restrict our attention, but sometimes we might stumble upon an important discovery that may not have been of primary interest to us. This is why we want to emphasize that during the earlier phases of research when we are trying to generate ideas, narrowing the focus of observation may be counterproductive, stifling perhaps our imagination and creativity. Here you might be better served by a frame of mind in which your attention and observation act not so much as a narrow spotlight of illumination of a particular area of interest but as a lantern casting a diffuse radiance over the panoply of experience (Lehrer, 2009). Such a prepared yet open mind increases the chances of accidentally discovering something fortunate. This is known as the serendipity effect. Lore also has it that the seeds of Isaac Newton’s law of universal gravitation can be traced to his observations of the famed apple falling out of a tree. The point is that serendipity teaches us the value of keeping our minds and hearts open for the unexpected, unlikely, and counterintuitive.

### Starting With Theory

In many instances, your research question will flow directly from a theory. For example, as we have learned, Charles Darwin was the among the first to raise the
question of whether physiological changes may lead to, if not very well be the cause of, an emotion as opposed to being just the consequence of that emotion. Echoing the sentiments of Darwin, the 19th-century psychologist William James in 1884 and Danish physiologist Carl Lange in 1887, each working independently, developed remarkably similar ideas as to the nature of human emotions. They each theorized that emotions are feelings that come about as a result of physiological changes, such as perspiration, rapid heartbeat, muscular tension, and dryness of mouth, created by the autonomic nervous system that regulates bodily reactions to stress (Lang, 1994). Known as the James–Lange theory, it addresses the chicken-and-egg question by stating that the physiological changes come first, followed by the experience of emotional feelings (Ellsworth, 1994). So, for example, according to this theory, you see a bear, your muscles tense, your heart races, you feel afraid. In the 1920s, Walter Cannon and Philip Bard would argue that the James-Lange theory got it backward, and in what became known as the Cannon–Bard theory of emotional experiences, these researchers proposed that emotions come first, followed by bodily changes.

As a rapprochement of these two theories, consider recent research by Niedenthal (2007) on the ways in which various emotional states can be encoded in particular bodily movements and postures. To illustrate her point, Niedenthal offers the following amusing “thought experiment”:

A man goes into a bar to tell a new joke. Two people are already in the bar. One is smiling and one is frowning. Who is more likely to “get” the punch line and appreciate his joke? Here is another: Two women walking over a bridge. One is afraid of heights, so her heart pounds and her hands tremble. The other is not afraid at all. On the other side of a bridge, they encounter a man. Which of the two women is more likely to believe that she has just met the man of her dreams? (p. 1002)

As Niedenthal wrote, in both these scenarios, the first person, as you would guess, is the correct answer. How is this so intuitively obvious to us? Each vignette involves processing of emotions—the smiling approachable patron or the woman’s fear of heights. When we perceive and think about their emotions, we experience or more precisely re-experience in ourselves those subtle physical and mental changes of the relevant emotions. This re-experiencing is referred to as the “embodiment” of emotion. It involves perceptual, somatovisceral (body), and motoric re-experiencing of the relevant emotion. The embodiment theory of emotion thus proposes a dynamic interplay of specific bodily states and their associated emotions. A schematic of the theory taken from Niedenthal is presented in Exhibit 2.3. On the left side, you see that the picture of a snarling bear triggers a cascade of events across the brain and somatovisceral systems. On the right side, simply thinking about the snarling bear is enough to re-engage parts of the original visual system, which in turn reactivates parts of the brain and somatovisceral organs originally triggered. As experimental support for this theory, Niedenthal cited studies that have shown that (a) when people adopt emotion-specific bodily postures, they report experiencing the associated emotions; (b) when people make facial expressions or emotional gestures, their
perceptions and impressions are affected; and (c) inhibiting people’s motor movements can interfere with their experience of emotion.

Thus, we see the genesis of a research question. We can trace the research questions of Paul Ekman to competing views of the James–Lange and Cannon–Bard theories of emotion, which gave way to the modern-day embodiment theory of emotion. We see how Ekman posed his questions in such a way that could be directly answered using the scientific method.

**Literature**

Scientific literature can also serve as a fountain for new ideas. These ideas may come to us when we consider empirical findings of various peer-reviewed studies that are published in scientific journals. These journals constitute the scientific literature. **Table 2.2** presents a partial list of major peer-reviewed publications in
psychology. In looking at this list, you see that the peer-reviewed publications fall into one of two categories: empirical articles and review articles. An empirical article reports on a particular study and is written in a certain format divided into sections with an Abstract, Introduction (which will contain a review of the relevant literature that is the focus of the study), Method, Results, and Discussion. A review article examines several studies of a particular phenomenon, such as facial expressions of emotion. It evaluates the methodology used across different studies, examines the degree to which findings are robust across various conditions, settings and procedures, and comments on the extent to which the empirical findings allow for general theoretical conclusions.

An additional potential wealth of inspiration for new ideas may be found in the popular science literature, often written by eminent scientists who aim to explain science for a general audience. An example of an outstanding work of

<table>
<thead>
<tr>
<th>Empirical Articles</th>
<th>Review Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior Therapy</td>
<td>American Psychologist</td>
</tr>
<tr>
<td>Child Development</td>
<td>Annual Review of Psychology</td>
</tr>
<tr>
<td>Cognitive Psychology</td>
<td>Psychological Bulletin</td>
</tr>
<tr>
<td>Developmental Psychology</td>
<td>Psychological Review</td>
</tr>
<tr>
<td>Journal of Abnormal Psychology Science</td>
<td></td>
</tr>
<tr>
<td>Journal of Applied Psychology</td>
<td></td>
</tr>
<tr>
<td>Journal of Educational Psychology</td>
<td></td>
</tr>
<tr>
<td>Journal of Experimental Child Psychology</td>
<td></td>
</tr>
<tr>
<td>Journal of Experimental Psychology: General</td>
<td></td>
</tr>
<tr>
<td>Journal of Experimental Psychology: Learning, Memory, and Cognition</td>
<td></td>
</tr>
<tr>
<td>Journal of Experimental Psychology: Human Perception and Performance</td>
<td></td>
</tr>
<tr>
<td>Journal of Personality and Social Psychology</td>
<td></td>
</tr>
<tr>
<td>Proceedings of the National Academy of Sciences</td>
<td></td>
</tr>
<tr>
<td>Psychological Assessment</td>
<td></td>
</tr>
<tr>
<td>Psychological Science</td>
<td></td>
</tr>
<tr>
<td>Neuropsychology</td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td></td>
</tr>
</tbody>
</table>
popular science is the book by renowned evolutionist Professor David Sloan Wilson of Binghamton University, *Evolution for Everyone: How Darwin’s Theory Can Change the Way We Think About Our Lives* (2007). Closely related to popular science is **science journalism** that often focuses on recent developments in science that are judged newsworthy. Malcolm Gladwell’s (2002) piece in the *New Yorker* on Paul Ekman is a classic example of science journalism. Both popular science and science journalism fulfill an important public service of bringing peer-reviewed publications borne of the scientific method into popular political and cultural discourse.

Pieces in both popular science and science journalism are often exceptionally well-written works of literary distinction: lively, clear, and engaging and free of technical jargon. The downside is that this popular literature lacks the critical sense of proportionality, cautiousness, and tentativeness that is fundamental to the scientific method. Scientific findings can be oversold, if not downright sensationalized. Michael Shermer, prolific science journalist and scholar, warns of such overselling in the area of brain scan studies, citing as an example of sensationalizing findings a November 11, 2007, *New York Times* opinion piece titled “This Is Your Brain on Politics.” by Marco Iacoboni and colleagues. Likewise, in a November 7, 2009, *New York Times Book Review*, eminent psychologist Steven Pinker criticized science journalist Malcolm Gladwell, opining “. . . when a writer’s education on a topic consists in interviewing an expert, he is apt to offer generalizations that are banal, obtuse or flat wrong.”

How then might you view this popular literature? We emphasize these works of journalism only as a source of ideas that might inspire and motivate you to pursue a formal line of investigation using the scientific method. These journalistic pieces represent important **secondary sources**, that is, second-hand media accounts of scientific work. They are best used as clues to lead you to the **primary source**, the first-hand empirical report published in a peer-reviewed journal.

Consider an interesting newspaper article by journalist Javier Zarracina that appeared in the April 4, 2010, *Boston Globe* titled, “A Life Lesson From the Majors. Smile, You’ll Live Longer.” The newspaper account describes an empirical study published in the peer-reviewed journal *Psychological Science* by researchers Ernest L. Abel and Michael Kruger (2010) titled “Smile Intensity in Photographs Predicts Longevity.” With a sample of 230 photos of baseball players culled from the Baseball Register for 1952, Abel and Kruger coded the intensity of players’ expressions as either *no smile*, *partial smile* (only movement of muscles around the mouth, that is, only contractions of the zygomatic major muscles), or *full smile* (movement of muscles around the mouth and corners of the eyes, that is, contraction of both zygomatic and obicularis oculi muscles). They also compiled life data for the baseball players and controlled for other longevity factors, such as education, marital status, and body mass index. Their results showed that players with full smiles lived the longest, to 79.9 years, which is nearly 2 years more than the typical life expectancy for an American. Even a partial smile added to longevity, with these players living for 74.9 years, on average, 2 full years more than those who exhibited no smile. Their findings demonstrated an overall gain of 7 years of life for
smiling players. In addition, by including ratings of each of the player’s headshot on a three-point attractiveness scale, the researchers effectively ruled out perceived attractiveness as an alternative explanation for their findings: Unlike a player’s smile, a player’s perceived attractiveness did not influence his longevity.

One of the Research in the News exercises at the end of this chapter features the Abel and Kruger (2010) study along with its media account. It provides a sound example of how you can go from a secondary source from a newspaper to a primary source of an empirical study published in a peer-reviewed journal. And lest you smirk and think after reading the newspaper account that the Abel and Kruger study is a fluke, you would soon learn otherwise. Indeed, you would learn from the primary source that their findings add to those of numerous other studies that have shown the face to be an intriguing barometer of emotions. These include studies that have linked marriage stability and satisfaction with smile intensity of childhood photos and college yearbook photos (Harker & Keltner, 2001; Hertenstein et al., 2009). Likewise, Todorov and colleagues (2005) showed that their research participants who made split-second judgments of competence from faces of political candidates predicted the outcomes of U.S. congressional elections better than chance (e.g., 68% of the Senate race in 2004).

**Searching the Literature**

The Internet has been described as a godsend for researchers and scholars, especially for those who like to work in their pajamas (Tuhus-Dubrow, 2008)! Tracking down a published peer-reviewed article on the Internet can often be done in a matter of seconds, all from the comfort of your home. In addition, many researchers have very accessible and highly informative websites that include articles that can be identified by typing their name into a search engine such as Google and then easily downloaded free of charge. Internet traveling, so to speak, from hyperlink to hyperlink may also lead to an online serendipity of stumbling onto an unexpected article that could add a new dimension to your thinking or perhaps lead you down a very different path from your original idea.

A branch of Google called Google Scholar (http://scholar.google.com) specializes in searches of the scholarly and scientific literature: Just type a researcher’s name, such as “Paul Ekman,” and within seconds you will have access to a corpus of scholarly and scientific work on facial expression of emotion. Two excellent websites for major psychological organizations are American Psychological Association (www.apa.org) and Association for Psychological Science (www.psychologicalsciences.org). PsycINFO and PsycLIT are specialized, noncommercial search engines sponsored by the American Psychological Association (APA) to serve as databases for citations and abstracts of the world’s literature in psychology and related fields. So too is the Social Sciences Citation Index (SSCI), which includes articles from not only psychology but also related fields, such as sociology and criminology. With SSCI, you can identify a “key article” on your topic. SSCI will then provide you with a list of articles that have cited your key article. This list will give you a start in developing a bibliography of articles relevant to your topic.
These websites are not free and are proprietary, open only to subscribers. For example, websites of scientific journals are generally proprietary, meaning that only paid subscribers can access their contents of peer-reviewed articles. Most college and university libraries have subscriptions to a number of proprietary websites that can be accessed through the library system. Consulting with the library staff of your college or university will help you learn how to access proprietary websites of scholarly information.

Let the Searcher Beware!

Elizabeth Kirk (2002) offers some excellent guidelines for evaluating information from the Internet that are available on the Johns Hopkins University web page: www.library.jhu.edu/researchhelp/general/evaluating. Kirk’s main take-home point is that you should never use information that cannot be verified by other independent sources. To make her point about the dangers of accepting Internet information as gospel, she cites the famous Latin phrase caveat lector: Let the reader beware. She lists several criteria against which all information should be evaluated. Among these are authority (e.g., Who wrote it?), publishing body (e.g., Who is the sponsor of the site?), point of view or bias (e.g., Does the website advocate a point of view or sell media products, such as books, pamphlets, videos, etc.?), connection to the scientific literature (e.g., Does it cite references to peer-reviewed articles?), verifiability and transparency (e.g., Is the methodology open to public scrutiny and able to be reproduced?), and timelines (e.g., What is the date, and does the information get updated?). Kirk concludes by citing Yeats: “Cast a cold eye . . . on everything you read. . . . Look for other sources that can authenticate or corroborate what you find. . . . Learn to be skeptical and then learn to trust your instincts. . . .” Last, keep in mind that general search engines such as the family of Google search tools can be very useful, but they are commercially based and derive a portion of their revenues through advertisement.

Researchers have now begun to study how the Internet may shape scholarship. In a 2008 article in the prestigious peer-reviewed journal Science, University of Chicago sociologist James Evans argues that the Internet may paradoxically have a narrowing effect on thinking and research, suggesting that good ideas may be ignored and lost. Other scholars, such as anthropologist Alex Bentley of Durham University in England, quoted in the Boston Sunday Globe of November 23, 2008 by journalist Rebecca Tuhus-Dubrow, have warned that the Internet “makes academic research a popularity contest” dominated by search engines that order or prioritize information in ways that may have little to do with scientific merit or impact. Further complicating matters is that some commercial search engines may sell spaces to advertisers; in turn, this may affect how these search engines browse for information, how they list information, and how they update information.

The common lesson is that the tremendous benefits of the Internet, especially in the early phases of generating ideas, may also come with some potential costs. As we know from Chapter 1, we are all prone to confirmatory bias, which may lead us to only search for information on the Internet that supports our positions and ideas. We should always remain cognizant that the Internet and the search tools
used to navigate the electronic world have limitations that may affect how we search and gather information in ways we may not fully appreciate.

**RESEARCH STRATEGIES**

In conducting research, we are attempting to connect theory with empirical data, that is, evidence we obtained through scientific studies. As we have learned, researchers may make this connection by starting with a psychological theory and testing some of its implications with data. This is the process of **deductive research**; it is most often the strategy used in experimental studies. Alternatively, researchers may develop a connection between psychological theory by first systematically collecting observations, measurements, or data and then developing a theory that explains the patterns in the data. This **inductive research** approach is most often used in nonexperimental studies, such as those that use naturalistic observation methods. As you’ll see, a research study can draw on both inductive and deductive strategies.

Thus, two of the most important elements of all scientific research strategies are (a) data and (b) theory. Data are the empirical observations that allow for evaluating a theory. A theory is a set of propositions that explains a variety of occurrences. A theory performs three major functions: organization, explanation, and prediction. A good theory is one that is parsimonious and precise as well as powerful in its breadth and depth of explanation. It can explain a variety of occurrences with the fewest theoretical assertions. **Exhibit 2.4** depicts the link between theory and data. As you can see, a deductive research strategy will begin with a theory from which particular statements are drawn and then tested by collecting data. By contrast, an inductive strategy first examines the data and then derives a theory to explain the patterns shown in the data.

**Inductive Research**

Inductive research entails reasoning from particular data or empirical observations to a general theory. In Chapter 1, we were introduced to different research methods and designs. Among those that are probably best suited for an inductive research approach is the naturalistic observation method. Often used in **qualitative research** (which will be discussed in Chapter 11), a

**Exhibit 2.4** The Links Between Theory and Data

naturalistic observation design studies people in their natural settings so that their behaviors and words can be put into their proper context. Such descriptive study of people is also sometimes referred to as ethnography. Here observation, it is important to emphasize, does not mean the casual “seeing” of everyday life that leads to haphazard impressions. To the contrary, for this research methodology to be effective, observation must be systematic, which means that it must be conducted carefully, with precise description that allows for consistent or reliable cataloging of data and the orderly classification and analysis of that information (Adler & Adler, 1994). While qualitative researchers use naturalistic observation methods that tend to avoid predetermining categories of action that can be precisely measured, they, like their quantitative counterparts, make sure their studies yield reliable and valid data. In short, the aim of qualitative research is to understand context—the what, how, when, and where of an event or an action. It yields data regarding meanings, concepts, definitions, characteristics, metaphors, symbols, and descriptions of events or actions (Berg, 2004). As such, it is ideally suited for an inductive research approach.

Inductive research also fits well with quantitative studies that use naturalistic observation. Consider a study by David Sloane Wilson and his then graduate student Kevin Kniffin (Kniffin & Wilson, 2004) published in the peer-reviewed journal Evolution and Human Behavior. Here you find a highly user-friendly, naturalistic observation methodology to investigate to what extent attractiveness of faces is influenced by physical traits or nonphysical features such as knowing, liking, and respecting the person. Kniffin and Wilson (2004) had a college student sample rate photographs of classmates they knew taken from their high school yearbooks for familiarity, liking, respect, and physical attractiveness. Then the researchers had the same photographs rated for physical attractiveness by another person of the same age and sex as the yearbook owner who did not know the people in the photographs. Kniffin and Wilson reasoned that if the ratings of physical attractiveness were based solely on physical traits observable in the photographs, then the two raters should be in agreement, regardless of what one knew of their nonphysical qualities of knowing, liking, or respecting the person in the photograph. Their simple study showed that it was not the physical qualities but rather the nonphysical qualities that had a pronounced influence on the assessment of physical attractiveness of these yearbook photographs. As Wilson (2007) wrote in Evolution for Everyone, “As the theory predicts, it is not how well you know a person (familiarity per se) but how well you regarded him or her (liking and respect) that makes the biggest difference” (p. 121). Wilson induced an evolutionary interpretation of the data from the yearbook study, theorizing that attractiveness is largely determined by nonphysical features, such as trustworthiness and respect, as knowing these qualities of a person is vital for adaptation, fitness, and reproduction. To underscore his point, Wilson (2007) presented in Evolution for Everyone the amusing anecdote of the iconic photograph of Abraham Lincoln as an example of a “person made beautiful by his nonphysical qualities” (p. 123). Wilson tells the story that people regarded Lincoln in his lifetime as hideously ugly. Lincoln, himself, Wilson recounted, even poked fun at his appearance, as when accused of being two-faced, Lincoln quipped, “If I had two faces,
think I’d be wearing this one?” Yet his nonphysical qualities embodied in his personal history, character, and accomplishments have made the pictorial image of Lincoln among the most admired, revered, and, yes, beautiful.

Deductive Research

In deductive research, reasoning proceeds from a general theory to particular data. Theories, however, are not directly testable. Thus, hypotheses must be derived from a theory. These hypotheses are then tested and the ensuing results are used to evaluate how well the theory works—that is, how well it describes, explains, and predicts the nature of the phenomenon under study. A hypothesis is defined as a specific testable statement that can be evaluated by empirical observations or data. It is different from generalizations that are embedded in a theory. A generalization is a broad statement that cannot be directly tested but rather needs to be translated into one or more hypotheses. That is, a theory will consist of a host of generalizations, and each generalization can lead to one or more hypotheses. Only hypotheses are directly testable. This is so because a hypothesis proposes a specific relationship between two or more variables that must be measurable. And as we know, one of these variables will be the independent variable that is hypothesized to predict, influence, or cause variation in the dependent variable. In deductive research, then, the investigators formulate one or more hypotheses and develop particular methods and procedures of measurement and/or experimentation. This yields data that serve as a direct test of a given hypothesis.

Hypotheses can be worded in several different ways, and identifying the independent and dependent variables is sometimes difficult. When in doubt, try to rephrase the hypothesis as an “if-then” statement: “If the independent variable increases (or decreases), then the dependent variable increases (or decreases).” Exhibit 2.6 (Examples of Hypotheses) presents several hypotheses with their independent and dependent variables and their “if-then” equivalents. Exhibit 2.6 demonstrates another feature of hypotheses: direction of association. When researchers hypothesize that one variable increases as the other variable increases, the direction of the association is positive (Hypotheses 1 and 2). When one variable decreases as the other variable decreases, the direction of association is also positive. But when one variable increases as the other decreases, or vice versa, the direction of the association is negative or inverse. Hypothesis 2 states a negative
or inverse association between increased age and reduced motor speed. Hypothesis 4 is a special case in which the independent variable is categorical (which we will discuss Chapter 4), as in this example, where students study for an exam, either in the library or in their dorms. This qualitative variable cannot be said to increase or decrease. In this case, the concept of direction of association does not apply, and the hypothesis simply states one category or level of the independent variable is associated with higher values of the dependent variable.

As we know from Chapter 1, both the strength and direction of the association or relationship between two variables can be measured statistically by the correlation coefficient (symbolized as $r$), which can range from .00 to +1.00 and .00 to –1.00. If there is no relationship between two variables—that is, variation in one variable has nothing to do with variation in the other variable—then the correlation value will be .00 or close to .00. The closer a correlation between two variables is to 1.00, either +1.00 or –1.00, the stronger is the relationship between the two variables. The positive and negative signs of the correlation coefficient tell the direction of the association between the two variables. A positive correlation coefficient (a “plus” sign) means that there is a positive linear relationship: As scores on one variable increase, so do scores on the other variable. A negative correlation coefficient indicated by a “minus” sign means that there is a negative linear relationship: As scores on one variable increase, scores on the other variable decrease.

<table>
<thead>
<tr>
<th>Original Hypothesis</th>
<th>Independent Variable</th>
<th>Dependent Variable</th>
<th>IF-THEN Hypothesis</th>
<th>Direction of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Deeper encoding of words, better memory</td>
<td>Levels of encoding</td>
<td>Recall of words</td>
<td>IF encoding level increases, THEN memory increases</td>
<td>+</td>
</tr>
<tr>
<td>2. Older age, reduced motor speed</td>
<td>Age</td>
<td>Motor speed</td>
<td>IF age increases, THEN, motor speed decreases</td>
<td>–</td>
</tr>
<tr>
<td>3. Higher SAT scores, better GPAs</td>
<td>SAT</td>
<td>GPA</td>
<td>IF SATs are higher, THEN GPAs are higher</td>
<td>+</td>
</tr>
<tr>
<td>4. Exam grades are higher for students who study in the library vs. dorm</td>
<td>Place of study</td>
<td>Exam grade</td>
<td>IF place of study is the library THEN exam grades are higher compared to those who study in dorm</td>
<td>NA</td>
</tr>
</tbody>
</table>

The Research Circle

In reality, the process of conducting research designed to test explanations for psychological phenomena involves a dynamic interplay of moving from theory to data and then back to theory. This process can be characterized as a research circle. As Exhibit 2.7 shows, deductive and inductive research processes can be closely intertwined. With deductive research, theory gives birth to hypotheses, which are then tested by data. With inductive research, data give birth to an empirical generalization—a statement that describes patterns found in the data, from which a theory is formulated. The goals of inductive and deductive research approaches are identical: to develop and formulate theories, a set of general propositions that serve to organize and interpret data or to generate predictions for events and actions for which no data have yet to be obtained.

SUBSTANCE OF PSYCHOLOGICAL RESEARCH

In this next section, we examine the substance of psychological research, the nuts and bolts of formulating a hypothesis drawn from theory, creating variables, devising an experimental task, and analyzing data. To learn about the substance of psychological research, we use a simple experiment by Strack and colleagues (Strack, Martin, & Stepper, 1998), published in the peer-reviewed journal *Journal of Personality and Social Psychology* (JPSP), titled “Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis.”

Exhibit 2.7 The Research Circle
We chose this study mainly because it is a rather straightforward experiment that we think would be fun for you to replicate. It also, of course, had direct bearing on the research question on smiling and emotion that began this chapter.

**Facial Expression and Emotion**

We know that a good theory generates testable hypotheses that can be evaluated empirically with the scientific method. Recall also that a testable hypothesis is framed as a statement often in the form of a prediction that is made prior to the actual collection of data. Strack and colleagues devised their experiment as a test of the facial feedback hypothesis. The facial feedback hypothesis originates from the James–Lange theory of emotion. In contemporary research, the facial feedback hypothesis falls under the embodiment theory of emotion (Niedenthal, 2007), which we discussed earlier. The idea is that facial movements by themselves can re-enact particular emotional feelings and that these emotional feelings can be created by simply “tricking” a person into making specific facial contractions. Thus, in the Strack and company study, some participants held a pencil between their teeth while performing a laboratory task that involved rating the funniness of different cartoons. Holding the pencil in the mouth this way covertly induced the individuals to smile—and the corresponding emotion state of happiness. Other participants were instructed to hold a pencil between their lips without touching the pencil with their teeth, and this covertly induced the individuals to frown. (see **Exhibit 2.8**). Strack and colleagues specifically hypothesized that participants who

**Exhibit 2.8** Participants in the Strack et al. (1988) study either held a pen between the lips to inhibit smiling, as in the left panel, or else held the pen between the teeth to facilitate smiling

*Source: © Chris Goldy.*
were led to smile would judge the cartoons as funnier than participants who were led to frown. You can see that this is an excellent example of hypothesis: It is (a) derived from the embodiment theory of emotion, (b) presented as an a priori prediction, and (c) eminently testable.

**Creating Independent and Dependent Variables**

You can also see this is a very simple but ingenious method to create an independent variable: One group of participants rated funny cartoons while holding a pen between their teeth (inducing a smile) whereas another group of participants rated the same cartoons while holding a pen between their lips without it touching their lips (inducing a frown). No fancy technology, just a familiar pen held in the mouth in different positions to create the independent variable. The independent variable had three conditions or **levels**, two of which were whether participants held a pen between either their teeth or lips. Every independent variable must have two or more levels. Why? Because you want to compare levels of an independent variable on a dependent variable, and thus you need at least two levels for comparison. Strack and company wanted to see whether perceived funniness of cartoons changed as a function of whether participants held a pen between either their teeth (inducing a smile) or their lips (inducing a frown).

In addition, they had a third condition or level in which a third group of participants rated the funniness of cartoons while holding a pen in their nondominant hand that they would not normally use for writing. Participants assigned to this nondominant hand condition formed the **control group** for the study, meaning that they did not make facial muscle contractions intended to induce either a smile or a frown. Thus, we can see that the Strack and company study created an independent variable with three levels of conditions: participants with the pen either held with their teeth (teeth condition) or lips (lip condition) or with their nondominant hand (nondominant condition). In addition, they used **random assignment** (see Chapter 1) in allotting participants to one of these three conditions. By using random assignment, as you recall from Chapter 1, these researchers ensured that each participant had an equal chance of being assigned to either the teeth, lips, or nondominant hand conditions.

All participants completed the same task: That is, they rated the funniness of different cartoons while holding a pen either held tightly between their lips (lips condition), gently between their teeth (teeth condition), or with their nondominant hand (nondominant condition). Note that the researchers labeled their independent variable as “condition” with three levels, but they could have easily called it something else such as “pen hold” with three levels—teeth, lips, and nondominant hand. A researcher has leeway in naming an independent variable, although simple and sensible labels are of course preferable. As you also know, the dependent variable is the observed effect, result, or outcome that is measured in response to a change in the independent variable. In the Strack and company study, the dependent variable was the humor ratings their research participants completed. The researcher also has leeway in labeling a dependent variable, but it makes little
or no sense to use arbitrary names. For example, there is no better name for the dependent variable in the Strack and company study than humor “ratings” of cartoons that they had their participants complete.

We learned earlier in the chapter about the difference between conceptual and operational definitions. Using this distinction here, we can think of both independent and dependent variables as falling under the category of operational definitions. The independent variable specifies the operations that will be used as a manipulation—for example, what participants will do or how they might be grouped. The dependent variable specifies the operations that will be used to measure the effects of the independent variable on performance, which can be assessed, for example, by ratings scores, observational recordings, physiological response, test scores, and the like. A **control variable** is a potential source of influence on the dependent variable that you as a researcher desire to hold constant. You are not interested in studying the effects of a control variable, but it often can be an unwanted source of influence that you need to assess. Recall from Chapter 1 that this unwanted source of influence can also be referred to as **confounds**, which can include so-called third variables. Thus, you aim to rule out any effect a control variable may have on the dependent and independent variables.

To recap, a researcher examines the effects of changes of an independent variable on a dependent variable while using control variables to rule out extraneous sources of influence (e.g., confounds, third variables). The independent variable is what you change, the dependent variable is what you observe, and the control variable is what you keep constant or quantify, if it can’t be held constant. Strack and colleagues, for example, held a variety of factors constant. These included a cover story, testing setting, and the testing stimuli (cartoons), meaning that all participants heard the same cover story, were tested in the same laboratory setting, and rated the same cartoons. These would be examples of control variables that Strack and colleagues held constant, as they did not want the setting or cartoons to affect the relationship of the independent variable of facial contractions on the dependent variable of humor ratings.

Unfortunately, though, there are countless other variables in addition to the independent variable that could have easily influenced the dependent variable. Indeed, it is impossible to have enough control variables to rule out all the possible extraneous, potentially confounding factors that could influence the relationship between the independent and dependent variables. These other factors could vary with or contaminate the relation of the independent and dependent variables. This, as we recall from Chapter 1, is a major problem of confounding variables, which occurs when one or more variables that are thought to be extraneous based on the theory under investigation are in fact related to both the independent and dependent variables. When this happens, you might want to scream “Damn!” as these confounding variables can pose alternative explanations for findings of a study.

For example, Strack and colleagues did not want the cognitive interpretation that people might make about their facial actions to influence their ratings of humor or to confound their results. That is why the researchers presented the cover story, which they hoped would ensure that their participants enacted the desired facial movements (i.e., smile vs. frown) without being aware of or recognizing the
corresponding emotion (i.e., happy vs. sad) or the real purpose of the study. Strack and colleagues also addressed the potential confounding problem of differences in difficulty among the three conditions. They did so by having their participants, while holding the pen in their mouth (lips vs. teeth), rate the degree of difficulty they experienced as they used the pen to complete their ratings. These ratings of difficulty served as a control variable, as the researchers did not want the degree of difficulty of the pen position to influence the humor ratings of cartoons. This allowed them to rule out different degrees of difficulty for the three experimental conditions as an alternative explanation for their findings.

**Evaluating With Statistics**

As we learned in Chapter 1, a testable hypothesis is evaluated statistically. If the statistical analyses of the data show that the obtained results are highly unlikely due to chance, then the hypothesis is accepted—the predicted relationship between, for example, two variables is considered to be highly likely. If, on the other hand, the statistical analyses of the data show that the obtained results are likely due to chance, then the hypothesis is rejected. Statistical probabilities are set at levels of significance in the form of probability (p)-values. Traditionally, p-values of .05 or less indicate that an obtained finding is statistically significant, meaning that there is a 5% chance or less that the obtained results are due to chance. Thus, *statistical significance* simply means that a finding is unlikely due to chance at a certain level of probability, usually .05 or less.

Now let us look at the results of the Strack and company study of 92 male and female undergraduate students at the University of Illinois (Study 1). Their hypothesis predicted that participants who were led to smile would rate the cartoons as funnier than participants who were led to frown. **Table 2.3** compares average or mean (M) funniness ratings as well as difficulty ratings for the participants who held the pen gently between the teeth, or tightly between the lips, or in the non-dominant hand. Ratings were made on a range of 0 to 9, with a lower value indicating lower funniness and difficulty, and higher value indicating higher funniness and difficulty. As predicted by the facial feedback hypothesis, participants induced to smile by holding the pen gently between their teeth gave the cartoons the highest funniness ratings (M = 5.14), followed by those who held the pen in their non-dominant hand (M = 4.77), and participants led to frown by holding the pen with their lips gave the cartoons the lowest funniness ratings (M = 4.32).

The million-dollar statistical question pertains to the different mean ratings among the three groups of participants. The question, put simply, is whether these differences are statistically significant or are simply a matter of chance. The good news for Strack and colleagues is that their statistical analyses indicated that the differences among the means were unlikely due to chance at a probability level of less than .05 (p < .05). This meant that the likelihood is less than 5 out of 100 that these differences were due to chance. Strack and company also needed to demonstrate that their humor rating results were not confounded by any differences in difficulty that their participants may have experienced in using a pen that they held either between their teeth, between their lips, or in their dominant hands.
Here, too, the news was good. As you can see in Table 2.3, even though participants led to smile evaluated the cartoons as the funniest, they also found writing with the pen held gently between their teeth to be the most difficult in relation to the other two groups. Thus, the researchers could correctly conclude that the differing degree of difficulty of holding the pen in either the teeth or lips did not confound how research participants rated the funniness of the cartoons.

<table>
<thead>
<tr>
<th>Table 2.3</th>
<th>Rating of Funniness and Difficulty</th>
<th>(Ratings ranged on a scale from 0 to 9, with lower ratings indicating lower funniness and difficulty and higher ratings indicating higher funniness and difficulty)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position of Pen</td>
<td>Cartoon</td>
<td>Lip</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>-----</td>
</tr>
<tr>
<td>First</td>
<td>3.90</td>
<td>5.13</td>
</tr>
<tr>
<td>Second</td>
<td>4.00</td>
<td>4.10</td>
</tr>
<tr>
<td>Third</td>
<td>4.47</td>
<td>4.67</td>
</tr>
<tr>
<td>Fourth</td>
<td>4.90</td>
<td>5.17</td>
</tr>
<tr>
<td>Mean funniness</td>
<td>4.32</td>
<td>4.77</td>
</tr>
<tr>
<td>Mean difficulty</td>
<td>4.47</td>
<td>2.72</td>
</tr>
</tbody>
</table>

research is to improve on previous studies, to assess, for example, the extent to which the findings are reproducible, replicable, or robust given different conditions, settings, and samples. How does one learn to evaluate research and to become a critical consumer of the science of psychology? In this section, we approach this problem from two related perspectives. We first present how to evaluate research from a general or macro perspective of an educated consumer. Moving from our lofty philosophical armchair, we then focus on specific and more technical aspects of a particular study or research report. Steeped in our burgeoning knowledge of the toolbox of methods and designs of psychological research, we present a step-by-step process for the professional reviewer.

**What Does It All Mean?**

Oftentimes, we will learn about interesting research second-hand through science journalism, such as a newspaper account or a television story. As we know, science journalism can be an excellent source for generating ideas. However, it is not subject to the peer-review process of a scientific journal. The great concern is that science journalism may not provide an accurate account of a study. If learning about a study from a secondary source, such as a newspaper article or a television show, keep in mind the goals of popular media are to attract attention, sell issues, and increase ratings; these outlets are not subject to the rigors of scientific peer review of professional journals, such as the *JPSP*, which published the Strack and company study. Thus, in order to critique a study, you must go to the primary source. In our example, you would read the Strack and colleagues study published in *JPSP*. You will find that the *JPSP* is a top-tier peer-reviewed periodical that carries the prestigious imprimatur of the American Psychological Association (APA). As we know, for a study to be published in an APA journal, it must past the muster of reliability and validity as evaluated through the rigors of peer review. So your first general criterion for evaluating a study is to consider the source: where an empirical report such as the Strack and company study is published. A credible source requires peer review.

Next, look to see if the goals of the study are clearly defined and articulated. Then consider whether the methods and design of the study are sufficient to achieve its goals. Ask yourself, “If I had to design this study to test this hypothesized relationship, what would I do?” In other words, you want to evaluate whether the study provides the best method for achieving the goals of the research. Here a mental template might be helpful: Begin with the broad distinction of whether you would classify a study as either experimental or nonexperimental. For example, Strack and colleagues aimed to examine whether facial muscle contractions could cause changes in emotional responses to cartoons. To test a causal relationship such as this requires an experiment of the kind carried out by Strack and colleagues. On the other hand, a correlational study, an important nonexperimental design in our proverbial research toolbox, does not allow you to infer causality. Indeed, a correlational study might demonstrate an important empirical relationship between two or more variables. However, a correlation does not imply causation, meaning that a relationship between two variables does not mean that changes in one cause changes in another.
Also, in developing an eye of a critical consumer of research, keep in mind our discussion of pseudoscience in Chapter 1. We know that sound research has a clear connection to previous studies. How does the study fit with previous research? If it is a truly novel discovery, what made it so? Did it correct a flaw that prevented other studies from demonstrating these results? Also, always keep in mind the difference between anecdotal and empirical evidence. Personal testimonial accounts can be very moving and convincing, but we know such anecdotes do not constitute evidence for the scientific method.

Last, look at who funded the research. Major federal funding agencies and private foundations fund only research proposals that have been evaluated carefully and ranked high by a panel of experts, or peer reviewers. The funding agencies also often monitor the progress of the research. This does not guarantee that every such project is good, but it goes a long way toward ensuring some worthwhile products. On the other hand, research that is funded by organizations that have a preference for a particular outcome should be given particularly close scrutiny (Locke, Silverman, & Spirduso, 1998).

Professional Review

Now let’s go back in time and imagine being asked by the editor of the Journal of Personality and Psychology to serve as a peer reviewer of the Strack and company study. Your task as a peer reviewer is to provide an opinion as to the scientific merit of the study. The manuscript in front of you will be organized into a specific format with specific sections: Abstract, Introduction, Method, Results, and Discussion. You carefully read the manuscript, and for each of these sections, there are particular questions to consider that will help you to critique the study and to offer an opinion as to its scientific merit.

Abstract

The abstract summarizes the major points of the study. As a reviewer, you want to learn from the abstract the problem under investigation, the procedures used to study the problem, the results, the conclusions, and the implications of the research findings. The Strack and company abstract provided a concise description of their study, including its rationale, methodology, results, and theoretical conclusions. From reading the abstract, you should understand the main finding or contribution of the study.

Introduction

The introduction identifies the problem to be investigated and why it is important. In reviewing the introduction, look to see if the problem and rationale have been cast in a psychological theory, which should serve as a framework for the study, defining its key concepts and terms. Two general questions are helpful guides for reading an introduction: What is the study about and why does it matter? More formal questions are: What is the theory and hypothesis and how are they related?
For both these sets of questions, clear answers can be easily inferred from reading the introduction of the Strack and company study.

That is, Strack and colleagues devised their experiment as a test of what is referred to in the literature as the facial feedback hypothesis. The facial feedback hypothesis, which originates from the James–Lange theory of emotion, states that facial movements by themselves can induce particular emotional feelings and that these emotional feelings can be created by simply “tricking” a person into making specific facial contractions. Strack and colleagues specifically hypothesized that participants who held a pen gently between their teeth would induce smiling, which in turn would lead to significantly higher ratings of the funniness of cartoons than would participants who held a pen tightly between their lips that presumably induced frowning. Thus, you can see that this is an excellent example of a hypothesis as it is derived from theory, presented as an a priori prediction, and eminently testable.

**Method**

The method section details the operations and procedures that the researcher used in the study. Two helpful hints in reading the method: First, the method section should provide sufficient detail so that another researcher could reproduce the study. Second, you as a reader should be able to put yourself in the shoes of a participant and understand exactly what the participant actually was instructed to do in the study. How participants were recruited and how the study sample was formed should be clearly stated in the method. In reading the method, look for any signs of sample bias that occur when participants are not randomly selected. Also, consider how representative is the sample, as this will affect the generalizability of the study results.

The method section should also make clear operational definitions for both the independent and dependent variables. What was the experimental task? Were participants randomly assigned to different levels of the independent variable? Now let’s see how the Stark and company method fared. We know that they recruited their participants from the student body at the University of Illinois. This would not be considered a random sample. Also keep in mind that a sample of college students has limited generalizability. So on these two criteria, the Strack and colleagues study came up a bit short. On the other hand, the experimental task was ingenious, especially well designed to test the hypothesis. The cover story was elaborate and convincing and very likely eliminated methodological problems of earlier studies by preventing participants from being aware of the emotional feeling associated with the particular facial muscle activity of smile versus frown. In addition, the cover story provided a means to minimize the problem of demand characteristics. Demand characteristics are cues available that tip off participants to the scientific hypothesis of the researcher and influence how participants respond to study stimuli. The researchers also collected difficulty ratings as a control variable in order to assess whether their results would be influenced by these extraneous factors. Finally, the method should include a clear statement on informed consent procedures, which we discuss in Chapter 3. No research should be done without informed consent of the participants.
Results

The results section presents the major statistical findings of the study. The results include both descriptive statistics and inferential statistics. Descriptive statistics summarize the data, usually in the form of averages and standard deviations (a statistical measure of variability). Inferential statistics in the form of tests of significance tell the probability of whether the observed differences found in the study were produced by random, or chance, factors. Also important for inferential statistics is an estimate of what is known as effect size. The effect size refers to the strength of the predicted or hypothesized relationship between the independent variable and dependent variable. For the Strack and company study, the effect size can be evaluated by looking at the difference in funniness ratings for the teeth condition ($M = 5.14$) versus the lip condition ($M = 4.32$). While these differences reached statistical significance, the effect size appeared rather modest, reflecting a difference of less than one funniness rating point.

Overall, Strack and colleagues provided clear descriptive data in the form of mean funniness ratings as well as mean difficulty ratings for each of their three conditions. They also used inferential statistics to determine that the differences in the funniness ratings were unlikely due to chance. With the exception of failing to provide any statistical estimate of effect size, the results section provided the reader with the necessary descriptive and inferential statistics that are critical to test the facial feedback hypothesis.

Discussion

In the discussion section, the major findings of the study are often restated but typically only in narrative form, and the theoretical meaning of the results is discussed. Important questions to consider in reading the discussion are: Do the empirical findings support the inferences and explanations? Does the discussion section duly express tentativeness and exercise cautiousness, which are essential to the scientific method? Are alternative interpretations carefully considered?

Perhaps the most important question for the discussion section is whether the limitations of the study are thoughtfully and explicitly discussed. These limitations may be related to the extent to which the operations and procedures adequately capture the intended phenomena. Is the experiment too unrealistic, too contrived—clever but lacking any real application to life? The extent to which a study or experiment approximates the actual real-life phenomenon under investigation refers to ecological validity. Related to but different from questions of ecological validity are those that pertain to the generalizability of the study and how representative or culturally diverse is the sample. These considerations fall under the general or broader category of external validity (ecological validity is one particular aspect of external validity). Strack and company used a college sample. They devised a simple yet ingenious experimental task. But a fair question to ask is whether the experimental task was too contrived. Probably not, but this potential limitation was not considered in the discussion. They also did not discuss the limitations of generalizability of their college student sample. On the other hand,
Strack and colleagues provided theoretical interpretations that were well supported by their empirical findings. Also on a positive note, their discussion raised the possibility of additional follow-up studies. Overall, the discussion could be improved by addressing more fully the limitations of the study. However, to be fair, the Strack and company study was written for experts in their particular area and published in a top-tier psychology journal that assumes a level of knowledge and experience in research well beyond what we have assumed in writing this textbook.

Literature Review

All psychological research presented in peer-reviewed journals generally begins with a literature review. Some psychology journals are exclusively devoted to publishing literature reviews of particular topics of interest. Consider, for example, what the prestigious *Psychological Bulletin* seeks to publish, as described on their web page:

Integrative reviews or research syntheses focus on empirical studies and seek to summarize past research by drawing overall conclusions from many separate investigations that address related or identical hypotheses. A research synthesis typically presents the authors’ assessments of the (a) state of knowledge concerning the relations of interest; (b) critical assessments of the strengths and weaknesses in past research; and (c) important issues that research has left unresolved, thereby directing future research so it can yield a maximum amount of new information. (http://www.apa.org/pubs/journals/bul/)

As you can see from this description, literature reviews serve an extremely important function in developing a scientific understanding of a topic. A sound literature review creates and evaluates a body of knowledge that is drawn from a synthesis of empirical findings across independent studies.

A critical reading of any literature review, however, should pay close attention to several factors. First, how were the empirical studies selected for review? A major limitation is that authors of reviews are relatively free to select studies and can choose on their own those empirical findings viewed as relevant. Two independent reviewers of the very same topic could reach very different conclusions! We know from Chapter 1 the problems of confirmatory bias. It is not hard to imagine how confirmatory bias may influence a reviewer to seek only empirical findings that fit with a certain school of thought. Second, most traditional reviews lack an objective benchmark to evaluate the strength of different empirical findings gleaned across various studies. How does one weigh the importance of one empirical finding over the other? Third, literature reviews are limited by publication bias. That is, reviews are based on studies that have been published. Studies often are unpublished because of negative findings, such as failing to replicate a certain phenomenon or effect. A review generally does not take into consideration unpublished studies, instead focusing exclusively on empirical findings that have been reported in peer-reviewed journals. Publication bias can be a particular problem.
for studies examining the effectiveness of a particular drug. In fact, all drug studies, often referred as “trials,” must now be filed with the Federal Drug Administration (FDA), including negative trials in which the drug failed to show the predicted effect. The FDA requires this in order to limit the effects of publication bias in its reviews of drug studies.

Because of these difficulties reviews now often incorporate meta-analysis as a method to provide a more unbiased evaluation of the literature. Meta-analysis is a statistical tool that provides an objective metric to weigh the strength of results from individual studies. The metric used provides an estimate of effect size (strength or magnitude of the demonstrated findings) for the results of each study included in the review. The effect size statistic allows for studies to be directly and objectively compared with each other. The average effect size across studies can be computed to provide an objective indicator of the overall strength of the empirical findings. This, in turn, might also provide an index of the degree generalizability of the empirical findings, as the review includes many studies varying in settings, samples, procedures, and time periods. Meta-analysis provides a statistical means to combat the biases of any reviewer.

CONCLUSIONS

Researchers aim to describe and explain thoughts, actions, and events in terms of a theory. However, a theory cannot be tested directly. Thus, researchers perform experimental and nonexperimental studies to test hypotheses that are derived from a theory. A dynamic interplay ensues, with data obtained through hypothesis testing acting as the lifeline for both building and evaluating a theory. A good theory, then, will be one that can account for a wide variety of data with the fewest explanatory concepts. A theory gains power when it can organize ideas and facts into a coherent pattern and predict the results of future research. Causality, though, will be difficult to establish for any theory in psychology, as even highly controlled experiments cannot rule out all potential confounding factors. Evidence for causality should establish (a) temporal precedence, (b) covariation of cause and effect, and (c) elimination of alternative explanations. The best benchmark remains the degree to which a theory can account for findings that are robust across various conditions, settings, and procedures.

Both inductive and deductive reasoning can figure prominently in theory formulation and explanatory research. We have likened this process to a research circle in which deductive reasoning goes from theory to hypotheses to data, whereas inductive reasoning goes from data to empirical generalizations to theory. Ideas, as the so-called grist for the mill for research aimed toward building scientific theory, can come from many sources. The most important of these is the psychological literature, which consists of studies that have met the rigors of peer review. Scientific journalism and popular science literature can be two other potential sources for ideas, but articles and reports appearing in these two venues are not subject to peer review. As such, these secondary sources can best serve as clues to
lead you to the primary source, the first-hand empirical report published in a peer-reviewed journal.

Finally, evaluating the scientific merit of research is a key skill. We hope you become an educated consumer of psychological research! Always consider the source, and in order to critique a study, go to the primary source. Keep in mind that correlation cannot prove causation and anecdotes are not evidence for the scientific method. When reading a primary source of an empirical article, evaluate whether the abstract summarizes the study, the introduction defines key concepts, the method specifies operations and procedures, the results present descriptive and inferential statistics, and the discussion covers interpretation of the findings as well as limitations of the research.

**END OF CHAPTER**

**Chapter Highlights**

- Researchers describe and define terms both conceptually and operationally, with the former providing the meaning, often rather broad in scope, of an abstract term, such as *intelligence*, *anxiety*, or *emotion*, and the latter specifying how a concept is coded, measured, or quantified.

- Researchers aim to develop theories that will provide parsimonious and precise explanation of the nature of the phenomenon under investigation.

- Because a theory cannot be tested directly, researchers perform experimental and nonexperimental studies to test hypotheses that are derived from a theory.

- Data obtained through hypothesis testing act as the lifeline for both building and evaluating a theory.

- A good theory is one that can account for a wide variety of data with the fewest explanatory concepts. A theory gains power when it can organize ideas and facts into a coherent pattern and predict the results of future research.

- Deductive and inductive reasoning can figure prominently in theory formation and explanatory research, which we have likened to a research circle in which deductive reasoning goes from theory to hypotheses to data, whereas inductive reasoning goes from data to empirical generalizations to theory.

- Ideas are generated from a variety of sources: observation, theory, and searching the literature, most notably peer-reviewed journals but also written works in science journalism and popular science.

- Empirical articles contain the following sections: Abstract, Introduction, Method, Results, Discussion, and References. The abstract summarizes the study, the introduction defines key concepts, the method specifies operations and procedures, the results present descriptive and inferential statistics, and the discussion covers interpretation of the findings as well as limitations of the research.
• A traditional literature review article evaluates the methodology used across different studies, examines the degree to which findings are robust across various conditions, settings, and procedures, and comments on the extent to which the empirical findings allow for general theoretical conclusions.

• Meta-analysis provides a more unbiased review of the literature than traditional literature reviews.

**Key Terms**

Abstract 54
Alternative explanation 34
Applied research 35
Autonomic nervous system 37
Basic research 35
Cannon–Bard theory of emotion 37
Conceptual definition 32
Confounding 34
Control group 49
Control variable 50
Covariation of the cause and effect 34
Deductive research 43
Demand characteristics 55
Descriptive statistics 56
Direction of association 45
Discussion 56
Ecological validity 56
Effect size 56
Embodiment theory of emotion 37
Empirical articles 38
Empirical generalization 47
Ethnography 44
Experimental 30
External validity 56
Facial feedback hypothesis 48
Generalization 45
Inductive research 43
Inferential statistics 56
Informed consent 55
Introduction 54
James–Lange theory 37
Journals 38
Levels 49
Meta-analysis 58
Method 55
Nonexperimental 30
Operational definition 33
Popular science 39
Primary source 40
Program evaluation 35
Proprietary 42
PsycINFO 41
PsycLIT 41
Publication bias 57
Qualitative research 43
Research circle 47
Results 56
Research in the News


2. In the *Time* January 16, 2009, article “How to Lift Your Mood? Try Smiling,” John Cloud tells the story how of being coaxed during his workouts by his personal trainer: “Relax your face.” Cloud writes that he remained skeptical about his trainer’s advice until he read a January 2009 *Journal of Personality and Social Psychology* study by Matsumoto and Willingham (2009) that compared facial expressions in two groups of judo athletes—blind and sighted competitors from the 2004 Olympics in Athens. This study prompted Cloud to discover one of the oldest questions in the study of emotion as to whether we all learn facial expression through our culture or are they genetically coded for everyone. By the end of his enjoyable piece, Cloud concludes, “... the emotional train does run in two directions between your brain, which may be screaming from the pain that your trainer is causing, and your face, which can—if you draw it into a relaxed expression—inform your brain that it shouldn’t be protesting so much. So next time you’re working out and grimacing push your facial muscles into submission. Look blank. You will find it’s easier to get through one more rep.”

What makes this piece an example of science journalism? How would you describe the reporting methods used in this article? How do the style, organization, and content of science journalism compare with the scientific method of peer-reviewed studies? Do you think Cloud’s conclusions are supported by the scientific literature?

Activity Questions

1. Researchers in psychology “play” with ideas that require conceptual and operational definitions, such topics as maximization, happiness, intelligence, anxiety, impulsivity, attention, and emotion. For each of these, provide conceptual and operational definitions.
2. Independent, dependent, and control variables also fall under the “must know” category. To understand them requires practice. Here are some study situations. For each of these, name the three kinds of variables.

a. You want to see if different instructional sets influence performance on a cognitive test. One instructional set presents the cognitive test as an ability-diagnostic measure of intelligence and the other instructional set presents the cognitive test as an ability-nondiagnostic laboratory measure of problem solving unrelated to intelligence.

Independent (what you change/manipulate) variable: instructional set
Dependent (what you observe/measure) variable: scores on cognitive test
Control (what you hold constant) variable: cognitive test, same test given under the two instructional sets

b. You want to see if facial expressions can influence perception. You decide to have subjects look at cartoons and rate them on a funniness scale either while holding a pen between their lips, which makes it impossible to contract either of the two smiling muscles, or while holding a pen clenched between their teeth, which has the opposite effect of making them smile.

Independent (what you change/manipulate) variable: facial expression induced by either holding a pen between lips or clenched between teeth
Dependent (what you observe/measure) variable: funniness scores/ratings
Control (what you hold constant) variable: cartoons, same whether viewed while holding a pen between lips or clenched between teeth; funniness scale

c. You been hired to evaluate the effectiveness of two diets: South Beach vs. Atkins. You want to make sure both dieting groups weigh about the same before starting either the South Beach or Atkins diet. So you weigh everyone before the diet and at the completion of the diet 6 weeks later.

Independent (what you change/manipulate) variable: diet (South Beach vs. Atkins)
Dependent (what you observe/measure) variable: difference in weight before and after
Control (what you hold constant) variable: prediet weight; length of time (6 weeks) on respective diets

d. You want to study the influence of Alzheimer’s disease on memory while controlling for age. What this means is that you will have two groups of people of about the same age; one group will be diagnosed with Alzheimer’s disease and the other group will be your control group of healthy persons.

Independent (what you change/manipulate) variable: group (Alzheimer’s disease vs. control)
Dependent (what you observe/measure) variable: scores on memory test
Control (what you hold constant) variable: age, memory test, same given to both groups
3. Transform each of the following statements or problems into at least two testable hypotheses.
   a. Talking on cell phones while driving should be against the law.
   b. Listening to Mozart will make people smarter.
   c. Steroids increase home runs in baseball.
   d. Eating a vegetarian diet increases your grade-point average.
   e. Money can buy you happiness.

4. In virtually all experiments, there is a degree of deception as researchers need to avoid creating bias by providing too much information before subjects have completed a study. What do you think about the ethics of deception in psychological research?

5. High school yearbooks can be used to study faces. Recall the Kniffin and Wilson (2004) study in which participants rated photographs of people they knew taken from their high school yearbooks. Then these same photographs were rated by another person of the same age and sex as the yearbook owner who did not know the people in the photographs. This would be a rather straightforward exercise to do in class, provided students have their high school yearbooks! Students rate the attractiveness of photographs of graduates they knew from their own yearbook, and then these same photographs are rated by another student of the same age and sex as the yearbook owner who does not know the people in the photographs. You can collect data and then combine them with ratings from other students who participated in the exercise. What are your independent and dependent variables? What is your hypothesis and why? What are your expected results and why?


   A man goes into a bar to tell a new joke. Two people are already in the bar. One is smiling and one is frowning. Who is more likely to “get” the punch line and appreciate his joke? (p. 1002)

   You probably could easily guess correctly that the smiling person would get the joke. But describe how the experimental study of Strack and colleagues could be used as evidence for your answer.

Review Questions

1. Why is a naturalistic observational methodology considered to be a nonexperimental study? What are the strengths and limitations of nonexperimental studies?

2. Describe how Paul Ekman’s studies of facial expression of emotions addressed cross-cultural differences.

3. Compare and contrast deductive and inductive research strategies.

4. Identify and explain the three criteria for establishing cause and effect.
5. Describe how a researcher develops and uses conceptual and operational definitions.

6. Describe how there are the three kinds of evidence that are needed to establish causality. Why should a psychological researcher always be skeptical about interpreting evidence of causality?

7. How might a third variable serve to confound the interpretation of results? What should a researcher do when concerned about third variables?

8. Explain how it is that a theory cannot be directly tested. If this is so, how does a researcher go about evaluating a theory?

9. When obtained results achieve statistical significance, what does that mean precisely?

10. External validity and ecological validity are two important considerations in evaluating research. What do they mean? How are they related? How are they different?

11. What considerations should you keep in mind when using popular search engines such as Google Scholar?

12. Science journalism and popular science represent two interesting sources for developing thoughts and ideas about research. How would you use these sources as leads in pursuing a research question?

13. How might confirmatory bias on the part of the researcher contaminate a literature review? What can be done to reduce confirmatory bias in literature reviews?

14. What is peer review and why is it important?

15. How would you review each of the sections of an empirical article?

16. What is the purpose of a control group?

17. Explain how an obtained result could achieve statistical significance but be of a small effect size.

18. Explain the fundamental truism that you cannot infer causality on the basis of correlation.

19. Explain how a correlation coefficient contains two critical pieces of information about the relationship of two variables: (a) the strength of the relationship and (b) the direction of the relationship.

20. How does publication bias influence a literature review?