

Flawed Self-Assessment

Implications for Health, Education, and the Workplace

David Dunning,¹ Chip Heath,² and Jerry M. Suls³

¹Department of Psychology, Cornell University; ²Graduate School of Business, Stanford University; and

³Department of Psychology, University of Iowa

SUMMARY—Research from numerous corners of psychological inquiry suggests that self-assessments of skill and character are often flawed in substantive and systematic ways. We review empirical findings on the imperfect nature of self-assessment and discuss implications for three real-world domains: health, education, and the workplace.

In general, people's self-views hold only a tenuous to modest relationship with their actual behavior and performance. The correlation between self-ratings of skill and actual performance in many domains is moderate to meager—indeed, at times, other people's predictions of a person's outcomes prove more accurate than that person's self-predictions. In addition, people overrate themselves. On average, people say that they are “above average” in skill (a conclusion that defies statistical possibility), overestimate the likelihood that they will engage in desirable behaviors and achieve favorable outcomes, furnish overly optimistic estimates of when they will complete future projects, and reach judgments with too much confidence. Several psychological processes conspire to produce flawed self-assessments.

Research focusing on health echoes these findings. People are unrealistically optimistic about their own health risks compared with those of other people. They also overestimate how distinctive their opinions and preferences (e.g., discomfort with alcohol) are among their peers—a misperception that can have a deleterious impact on their health. Unable to anticipate how they would respond to emotion-laden situations, they mispredict the preferences of patients when asked to step in and make treatment decisions for them. Guided by mistaken but seemingly plau-

sible theories of health and disease, people misdiagnose themselves—a phenomenon that can have severe consequences for their health and longevity.

Similarly, research in education finds that students' assessments of their performance tend to agree only moderately with those of their teachers and mentors. Students seem largely unable to assess how well or poorly they have comprehended material they have just read. They also tend to be overconfident in newly learned skills, at times because the common educational practice of massed training appears to promote rapid acquisition of skill—as well as self-confidence—but not necessarily the retention of skill. Several interventions, however, can be introduced to prompt students to evaluate their skill and learning more accurately.

In the workplace, flawed self-assessments arise all the way up the corporate ladder. Employees tend to overestimate their skill, making it difficult to give meaningful feedback. CEOs also display overconfidence in their judgments, particularly when stepping into new markets or novel projects—for example, proposing acquisitions that hurt, rather than help, the price of their company's stock. We discuss several interventions aimed at circumventing the consequences of such flawed assessments; these include training people to routinely make cognitive repairs correcting for biased self-assessments and requiring people to justify their decisions in front of their peers.

The act of self-assessment is an intrinsically difficult task, and we enumerate several obstacles that prevent people from reaching truthful self-impressions. We also propose that researchers and practitioners should recognize self-assessment as a coherent and unified area of study spanning many subdisciplines of psychology and beyond. Finally, we suggest that policymakers and other people who makes real-world assessments should be wary of self-assessments of skill, expertise, and knowledge, and should consider ways of repairing self-assessments that may be flawed.

Address correspondence to David Dunning, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853-7601, e-mail: dad6@cornell.edu; to Chip Heath, e-mail: chip.heath@stanford.edu; or to Jerry Suls, e-mail: jerry-suls@uiowa.edu.

There are three things extremely hard: steel, a diamond, and to know one's self.

—Benjamin Franklin, *Poor Richard's Improved Almanack* (1750)

Over their lifetime, people base thousands of decisions on impressions of their skill, knowledge, expertise, talent, personality, and moral character. A teenage violinist applies to music school on the basis of some notion she holds of her musical virtuosity. A college student decides against a career in science because he believes math is a beast he would never slay. A military officer volunteers to command a dangerous mission because he has confidence in his own bravery, leadership, and grace under pressure. An elderly patient dismisses her doctor's suggestions because she thinks she knows best about her health. Self-assessments of skill and character play a sizable role along the gamut of choices ranging from small, such as which restaurants to patronize (Setterlund & Niedenthal, 1993), to large, such as what college majors to choose and which careers to pursue (Fouad, Smith, & Zao, 2002; Lent, Brown, & Hackett, 1994).

Thus, whether people decide well in life depends, at least in part, on whether their self-assessments are accurate, that is, on how successfully they follow the classical admonition from the Delphic oracles to "know thyself." To the degree that people judge themselves accurately, they make decisions, big and small, that lead to better lives. However, to the extent that people misjudge themselves, they may suffer costly consequences by pursuing wrong paths and missing opportunities to take advantage of special skills and resources they truly own. Sometimes, self-misjudgments may involve only lost time or effort. The would-be Broadway singer who fails to understand that a string of failed auditions provides a fair indication of his (lack of) singing talent suffers only a few months of misspent youth, with perhaps no long-run consequence other than wistful memories of youthful nearness to fame. At other times, the consequences of flawed self-assessment can be more severe, as in the case of a novice airplane pilot who thinks he can take off into the fog without his flight instructor's supervision. Moreover, such consequences are not constrained to the self. A doctor too assured of her expertise at diagnosing chest complaints exposes her patients to risks that might be life-threatening.

In this monograph, we review basic scientific research on the accuracy of self-assessment. In doing so, we assert that self-assessments of skill and character tend to be much more imperfect than people suspect. People fall prey to biases that leave their self-assessments flawed in systematic ways that carry significant implications. To be sure, we do not argue that self-judgments are valueless. However, we do argue that people's capacity to evaluate themselves and predict their behavior is usually quite modest and often much more meager than common intuition would lead one to believe.

In making this argument, we wish to provide evidence for two observations. First, the forces that influence social behavior, including self-behavior, are complex—and people rarely have all the information they need to render accurate self-judgments.

Therefore, achieving accurate self-knowledge is an inherently difficult task, as Benjamin Franklin sagely noted in the quotation with which we opened this review. Second, even when people do have in hand certain types of information that would lead them to more accurate self-assessments, they tend to neglect this information, which leads them to worse assessments than they are capable of. Thus, people in many substantive and consequential circumstances hold opinions of themselves that cannot withstand objective scrutiny (see Dunning, 2005, for a more extensive discussion).

In this monograph, we review data showing that people's perceptions of their skills, knowledge, personality, and character often do not mesh with objective reality. These misjudgments can take on two different forms. First, people's general evaluations of their skills and character—such as whether they are good leaders or verbally skilled—tend not to be tethered very tightly to objective performances in tasks that should reflect those skills and character traits. Second, when people offer specific predictions about how they will behave in a particular future situation, they make predictions that differ systematically from their actual behavior when that situation arrives.

Next, we turn our attention to psychological processes that produce flawed self-perception. Flawed perceptions may arise in various ways, so we provide no overarching theory of error in self-judgment, but we review major themes and variations uncovered in recent psychological research that explain, at least in part, why people tend to hold opinions of themselves that diverge from objective reality.

We then turn to three major domains of social life—health, education, and work—to examine how these themes play out for people confronting tasks central to their everyday lives. In this discussion, we examine the extent to which there is continuity between the findings of laboratory studies and judgmental patterns found in the real world. We also examine the real-world costs, and perhaps benefits, of erroneous self-judgment.

Such an investigation integrating lab findings with real-world behavior and consequences is valuable for two reasons. First, over the past 30 years, a growing body of evidence has shown that people fall victim to a number of systematic judgmental errors in laboratory settings (for reviews, see Gilovich, Griffin, & Kahneman, 2002; Hastie & Dawes, 2001; Nisbett & Ross, 1980). This research tradition, however, has not been without critics, who question whether such laboratory errors transfer to real-world contexts, arguing that these studies often involve artificial and unfamiliar stimuli that are not representative of what people confront in their everyday lives (Funder, 1987; Gigerenzer, 1996; Gigerenzer, Hoffrage, & Kleinboelting, 1991). If erroneous self-judgments found among college students in a psychological laboratory are also revealed, for instance, among doctors and CEOs—and if these errors are found to have real-world consequences—then the import of the laboratory research would be more firmly established.

Second, researchers may observe judgmental errors in their laboratory, but they sometimes lag in finding ways to prevent or

circumvent those errors. In the real world, where consequences for error are greater, individuals and organizations may have already found ways to begin to alleviate the problem of flawed self-assessment. Thus, as we examine self-judgment in real-world contexts, we also note instances in which individuals and institutions work to correct or repair such flaws.

EMPIRICAL EVIDENCE ON FLAWS IN SELF-ASSESSMENT

Over decades of research involving a wide variety of domains and circumstances, psychologists have examined how accurately people judge themselves. The usual finding is that people have a modest level of insight, at best, into their skill and character. First, researchers tend to find fairly small correlations between perceptions of skill and objective performance. Second, people tend to be too optimistic about their talents, expertise, and future prospects.

Correlations Between Perception and Reality

When researchers correlate self-assessments of knowledge and skill against objective performance, the relationship they find is rarely strong. Typically, it is modest to meager, and sometimes it is null. For example, people's views of their intelligence tend to correlate roughly .2 to .3 with their performance on intelligence tests and other academic tasks (Hansford & Hattie, 1982).¹ Students' ratings of their academic skill during their first year of college correlate only .35 with the evaluations their instructors give them (Chemers, Hu, & Garcia, 2001). People's beliefs about their ability to detect lying among others correlate only .04 with their performance (DePaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997). In the workplace, the correlation between how people expect to perform and how they actually perform hovers around .20 for complex tasks (Stajkovic & Luchins, 1998).

It is in the health domain that divergences between self-perceptions of knowledge and reality have been most commonly documented. Adolescent boys' confidence in their knowledge about how to use condoms correlates only slightly with their actual knowledge (Crosby & Yarber, 2001). Family practice residents' self-rated skill at interviewing patients and soliciting relevant health information correlates roughly .30 with ratings provided by their instructors (Stuart, Goldstein, & Snope, 1980). The confidence of nurses in their knowledge of basic life-support tasks fails to correlate at all with their actual level of knowledge

(Marteau, Johnston, Wynne, & Evans, 1989). Physicians' self-rated knowledge about thyroid disorders also fails to correlate with their performance on a quiz on the topic (Tracey, Arroll, Richmond, & Barham, 1997). Perhaps the most sobering finding is that surgical residents' views of their surgical skill also fail to correlate with their performance on a standardized board exam (Risucci, Torolani, & Ward, 1989).

Meta-Analytic Evidence

These demonstrations are not isolated instances. In 1982, Mabe and West searched the literature for studies that had examined the relationship between self-perceptions of knowledge and objective performance and then analyzed the results of all these studies as a group (i.e., performed a meta-analysis). They observed a large range of correlations between self-perception and performance, but when they took the average of all these correlations, they found that self-perceptions correlated with objective performance roughly .29—a correlation that is hardly useless, but still far from perfection.² Some domains produced higher correlations than others. In athletics, where feedback tends to be constant, immediate, and objective, the typical correlation was .47. In the realm of complex social skills, where feedback might be occasional and is often delayed and ambiguous, it tended to be much lower (e.g., .04 for managerial competence and .17 for interpersonal skills). Other meta-analyses, one on the self-perceptions of students in the classroom (Falchikov & Boud, 1989) and one on the perceptions of employees in the workplace (Harris & Schaubroeck, 1988), found similar modest relationships between what people believe about their skills and the performances they achieve.

Self Versus Peer Assessment

An additional finding buttresses the conclusion that self-insight about skills and knowledge is modest. Complete strangers armed only with scant information about an individual can predict that person's skills and abilities almost as well as he or she can, despite the fact that the individual has a lifetime of self-information to draw upon. Borkenau and Liebler (1993) showed participants videotapes in which target individuals walked into a room, sat behind a table, read a standard weather report, and then walked back out of the room—actions that typically took 90 seconds to complete. Participants who viewed these tapes—and who had no additional information—provided ratings of intelligence that predicted the targets' scores on standard IQ tests almost as well as the targets' self-ratings. Similarly, Epley and Dunning (2004) asked college students to rate their current romantic relationship along five dimensions and to answer three quick questions about

¹One way to assess the strength of these correlations in less technical terms is this: Suppose that among 100 people, 50 describe themselves as "above average," and 50 describe themselves as "below average." Now suppose that the correlation between perception and reality is zero. If that is the case, then 50% of these people will be right in their self-categorization, which is the same accuracy rate they would achieve if they judged themselves by flipping a coin. A correlation of .20 between self-categorization of "above" and "below" average and the reality would raise this accuracy rate to 60% (with 40% misclassifying themselves); a correlation of .40 would raise it to 70% (with 30% misclassifying). At a perfect correlation of 1.0, accuracy would be 100%.

²Mabe and West (1982) noted that the low correlation between perception and reality may be due, in part, to error in the measurement of self-evaluations and objective performance; that is, the measures used in the correlation may be unreliable to some extent. Thus, Mabe and West calculated what the perception-reality correlation would be if both variables were measured perfectly, without error. They estimated that the correlation would rise to .42—better, but still moderate in magnitude.

the relationship (e.g., what hobby or activity do you most enjoy doing together?). Complete strangers reading this minimal information were just as accurate as the students themselves at predicting whether their relationship would still be intact 6 months later.

A person's acquaintances may actually predict that person's abilities and performance better than the person him- or herself. To be sure, people may be more accurate than their peers in predicting some of their more mundane behaviors (Shrauger, Ram, Greninger, & Mariano, 1996) and also their emotional reactions to events (Spain, Eaton, & Funder, 2000), but this advantage can reverse (Kolar, Funder, & Colvin, 1996), particularly for more consequential behaviors. Although the self-views of surgical residents do not predict their performance on standardized board exams, their supervisor's ratings do, as do the ratings of their peers who are equally inexperienced (Risucci et al., 1989). Peer ratings of leadership, rather than self-ratings, predict which naval officers will be recommended for early promotion (Bass & Yammarino, 1991). College students predict the longevity of their roommates' romantic relationships better than they do their own (MacDonald & Ross, 1999).

Unrealistic Optimism

Self-assessments turn out to be flawed in another way. People overestimate themselves. They hold overinflated views of their expertise, skill, and character. That is, when one compares what people say about themselves against objective markers, or even against what might be possible, one finds that the claims people make about themselves are too good to be true. This bias toward undue optimism, self-aggrandizement, and overconfidence is exhibited in many ways.

Above-Average Effects

People, on average, tend to believe themselves to be above average—a view that violates the simple tenets of mathematics. In a survey of nearly one million high school seniors, 70% stated that they had “above average” leadership skills, but only 2% felt their leadership skills were “below average.” On their ability to get along with others, almost all respondents rated themselves as at least average—with 60% rating themselves in the top 10% of this ability and 25% rating themselves in the top 1% (College Board, 1976–1977). College students think they are more likely than their peers to live past 80 and have a good job; they think they are less likely to acquire a drinking problem or suffer a heart attack (Weinstein, 1980).

Such above-average effects, as they are called, are not constrained to college students. Motorcyclists believe they are less likely to cause an accident than is the typical biker (Rutter, Quine, & Albery, 1998). Business leaders believe their company is more likely to succeed than is the average firm in their industry (Cooper, Woo, & Dunkelberg, 1988; Larwood & Whittaker, 1977). People think they are less susceptible to the flu than their

contemporaries, and as a result avoid getting flu shots (Larwood, 1978). Of college professors, 94% say they do above-average work (Cross, 1977). People signing up to bungee jump believe they are more likely to avoid injury than the average bungee jumper, although their friends and family do not share this impression (Middleton, Harris, & Surman, 1996). Ironically, people even state that they are more likely than their peers to provide accurate self-assessments that are uncontaminated by bias (Friedrich, 1996; Pronin, Lin, & Ross, 2002).³

Overestimation of the Likelihood of Desirable Events

People overestimate their ability to bring about personally desirable events. Lawyers overestimate the likelihood that they will win the cases they are about to try (Loftus & Wagenaar, 1988). Stock pickers think the stocks they buy are more likely to end up winners than those of the average investor (Odean, 1998).

People also overestimate the likelihood that their own future actions will be socially desirable, even though their predictions regarding their peers' behavior turn out to be more accurate. In one large lecture class at Cornell University, 83% of students predicted that they themselves would buy flowers in an annual charity drive for the American Cancer Society, but that only 55% of their fellow students would do the same. The actual percentage buying flowers 4 weeks later was 43% (Epley & Dunning, 2000). In another class, 90% claimed that they would vote in an upcoming presidential election, but that only 75% of their peers would. The actual percentage was 69% (Epley & Dunning, 2004). Other work has shown similar errors in self-prediction (and rough accuracy for peer prediction) when students predict how much money they will donate to charity, whether they would volunteer for a long experiment so that a little girl can participate in a short one, and whether their romantic relationships will last longer than half a year (Epley & Dunning, 2000, 2004).

Underestimation of Task-Completion Times

People also consistently overestimate how easily they can complete tasks (as measured by time or money), a phenomenon known as the *planning fallacy* (Buehler, Griffin, & Ross, 1994). For example, the amount of time college students take to finish their senior thesis is 3 weeks longer than their most “realistic” estimate of how long it will take—and 1 week longer than what they describe as their “worst case” scenario (Buehler et al., 1994).

³We should mention one important caveat about the generality of the above-average effect. Although Americans, Canadians, and Western Europeans tend to see themselves as above average along any desirable dimension, respondents from Far East Asia do not (e.g., Heine & Lehman, 1995; Heine, Lehman, Markus, & Kitayama, 1999). Thus, the generality of the above-average effect across cultures is not a given, and its exact cultural and geographical boundaries could be explored further. In addition, further studies could explore the generality of other biases described in this review. One should not assume that the presence of one bias in a culture implies the presence all other biases in that culture. Similarly, the absence of one bias in the culture does not imply the absence of all. For example, although Eastern respondents avoid the above-average bias, some tend to be more overconfident than Americans and Canadians in the judgments they reach (e.g., Yates, Lee, & Bush, 1997; Yates, Lee, Shinotsuka, Patalano, & Sieck, 1998).

In one illustrative study, researchers asked students who were working on a class assignment to indicate the time within which they were 50% certain they could finish the project, as well as the time within which they were 99% certain they could finish it. On average, if the students were accurate, about half would have finished by the 50% deadline and 99% would have finished by their very conservative 99% deadline. However, only 13% had actually finished by their 50% deadline, and only 45% had finished by their 99% deadline. Thus, even for a deadline that students were virtually certain they would meet, their confidence far exceeded their accomplishments (Buehler, Griffin, & Ross, 2002). In a similar vein, citizens typically believe they will complete their tax returns more than a week sooner than they actually do (Buehler, Griffin, & MacDonald, 1997).

Overconfidence in Judgment and Prediction

Finally, people place too much confidence in the insightfulness of their judgments, overestimating the chances that their decisions about the present are sound and that their predictions about the future will prove correct. This phenomenon is known as the *overconfidence effect*. College students overestimate the probability that their answers to general knowledge questions are correct (Fischhoff, Slovic, & Lichtenstein, 1977). They are also overconfident in their forecasts of what events they will experience over the course of a semester (Dunning & Story, 1991; Vallone, Griffin, Lin, & Ross, 1990), as well as in their forecasts of the events their college acquaintances will experience (Dunning, Griffin, Milojkovic, & Ross, 1990). Analysts at the U.S. Central Intelligence Agency overestimate the accuracy of their predictions about future world events (Cambridge & Shreckengost, 1980). Surgical trainees place too much confidence in their diagnoses after looking at X-ray evidence (Oksam, Kingma, & Klasen, 2000). After looking over a client's case materials, clinical psychologists overestimate the chance that their predictions will prove accurate (Oskamp, 1965).

Indeed, even when people are the most confident, that certainty is no guarantee of accuracy. In studies in which college students expressed absolute (100%) certainty in their answers, they still were wrong roughly one time out of every five (Fischhoff et al., 1977). In another study, when doctors diagnosed their patients as having pneumonia, predictions made with 88% confidence turned out to be right only 20% of the time (Christensen-Szalanski & Bushyhead, 1981).

PSYCHOLOGICAL MECHANISMS

A wide variety of psychological mechanisms underlie these flawed self-assessments, and it would be difficult, if not impossible, to catalogue all of them in a single document. Instead, we focus on two of the most widely documented biases described in the preceding section—above-average effects and overestimation of the likelihood of desirable events—and describe two general themes that explain why these biases arise. The first

theme is that people typically do not possess all the information required to reach perfectly accurate self-assessments. Gaining an accurate view of self is an intrinsically difficult task (Dunning, 2005). There are too many factors that are unknown, unknowable, and undefinable for people to make adequate evaluations of their performance or accurate forecasts about how they will act in the future. We certainly do not blame people for failing to know everything, but we can say that people should take into account what they fail to know and adjust their predictions accordingly. The second theme we discuss is that even when people do have valuable information that would guide them toward appropriate self-evaluations, they often neglect it or give it too little weight; thus, they make potentially avoidable errors.

Let us consider the above-average effect and the overly optimistic prediction of desirable events, in turn, to see how lack of information on the one hand and neglect of valuable information on the other serve to produce each effect.

Explanations for the Above-Average Effect

People lack crucial information they need when they compare themselves against others; they also ignore valuable information that they actually possess or could seek out. These twin themes are quite evident when one examines research on the above-average effect. We cite four informational deficits that lead people to believe they are doing much better than their peers. In addition, people neglect important information that could prompt them to reach more accurate conclusions.

Information Deficits

The Double Curse of Incompetence. People often do not have the knowledge and expertise necessary to assess their competence adequately. Consider, for example, the plight of the incompetent, who are often not in a position to recognize just how poor their decisions are. In many significant social and intellectual domains, the skills necessary to recognize competence are extremely close if not identical to those needed to produce competent responses. For example, recognizing whether an argument is logically sound requires a firm grasp of the rules of logic. If people do not understand the rules of logic, not only will they make logical errors, but they will also not recognize that their arguments are logically defective—or that anyone else's argument is logically superior. Thus, incompetent individuals suffer a double curse: Their deficits cause them to make errors and also prevent them from gaining insight into their errors.

Several studies have now shown that incompetent individuals (i.e., those performing poorly relative to their peers) fail to show much insight into just how deficient their performance is (Kruger & Dunning, 1999). College students scoring in the bottom 25% on a course exam walked out of the exam room thinking that they outperformed a majority of their peers (Dunning, Johnson, Ehrlinger, & Kruger, 2003). Debate teams performing in the bottom 25% at a regional tournament believed they were winning 59% of

their matches when in fact they were winning only 22% (Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2004). Medical students mishandling a mock interview with a patient rated their interviewing skills much higher than their instructors did (Hodges, Regehr, & Martin, 2001). In a hospital, lab technicians in the bottom 25% among their peers failed to realize that their performance was so low (Haun, Zeringue, Leach, & Foley, 2000).⁴

Other work also demonstrates that poor performers, relative to their more competent peers, have more difficulty differentiating accurate from inaccurate performance. Compared with good students, poor students less successfully identify which specific questions they have gotten right on an exam and which they have gotten wrong (Sinkavich, 1995). Novice bridge players are less likely than expert players to tell good moves from bad ones (Keren, 1987). Students with little experience in physics, compared with more accomplished physics students, have less accurate intuitions about which physics problems are generally difficult to solve (Chi, Glaser, & Rees, 1982). Paradoxically, although training people on logic improves their skill, such training also reveals to them past flaws in their logical reasoning, leading them to provide more pessimistic views of their logical reasoning ability at the moment their skill level rises (Kruger & Dunning, 1999).

Unknown Errors of Omission. But it is not only the incompetent who often do not have all the information they need to appraise their skill and performance accurately. For example, although people trying to solve a problem may find it easy to consciously critique the solutions they generate, by definition they are not aware of solutions they could generate but miss, that is, their *errors of omission*. For example, suppose we asked you to list as many English words as you could from the letters in the word *spontaneous* (e.g., *tan*, *neon*, *pants*), and you found 50. Whether this performance is good or bad depends, in part, on how many words are possible, and it is difficult to expect that you—or anyone else—would have an accurate intuition of what that figure is. In fact, more than 1,300 English words can be created from the letters in *spontaneous*.

Recent work shows that people tend to have little insight into their errors of omission (Caputo & Dunning, in press); however, they give these errors a good deal of weight (indeed, equal to what they give to the solutions they generate themselves) once they find out about them. For example, in one study (Caputo & Dunning, in press, Study 4), graduate students were given brief descriptions of research studies and asked to list all the methodological difficulties they could find. Students' initial evaluations of their knowledge of research methodology were not correlated with their objective performance on this task. But students provided

more pessimistic, and accurate, assessments of their knowledge about research methodology once their errors of omission (i.e., the study flaws they had failed to identify) were made known to them.

Uncertain Lessons From Feedback. People also receive incomplete feedback about their actions, which can lead them to harbor inflated views about the wisdom of their actions. Suppose an office manager takes a poorly performing employee aside and berates him. Next day, that employee performs better—a result that presumably provides evidence for the sagacity of the office manager's intervention. However, the manager does not know what might have been achieved by other alternatives, such as sitting down with the employee for a sympathetic talk, or even doing nothing. Perhaps these alternatives would have worked as well, or even better, but the office manager will never know (for further discussions about problems with real-world feedback, see Dawes, 1988, and Dunning, 2005).

The Ill-Defined Nature of Competence. Perhaps the most fundamental reason for people to have incomplete knowledge of their competence is that in many domains, what it takes to succeed is hard to define. It is easy, for example, to define successful math performance. In math, there are specific right answers that come with well-delineated algorithms designed to produce them. However, success in many domains is ill defined (Newell, 1969; H.A. Simon, 1973). No one has the exact optimal algorithm for composing a classic symphony, writing the great American novel, or generating an effective marketing plan—nor any way to know when someone has done the best job possible.

Many skills and personality traits themselves are ill defined in that many different criteria are arguably relevant for them. For example, how would one determine whether someone is the most intelligent person in the room? Does having a large vocabulary matter? What about ability to read other people's nonverbal behavior? Facility at math? Knowledge of fine wines? Recent research shows that people tend to take ill-defined traits and define them to their advantage. A person skilled in math considers math skill to be more centrally related to intelligence than does a person who is math challenged. A person who accepts social norms considers that habit a better sign of intelligence than does a nonconformist (Dunning, Perie, & Story, 1991).

As a consequence, people tend to believe themselves to be above average on traits that are ill defined, but not on ones whose definition is more constrained. For example, people tend to say they are more sophisticated, idealistic, and disciplined than their peers (ambiguous traits all), but are not likely to think they are any more neat, athletic, and punctual (traits that are more constrained in their meaning; Dunning, Meyerowitz, & Holzberg, 1989; Suls, Lemos, & Stewart, 2003). On ill-defined desirable traits, college students rate themselves more favorably than their roommates rate them (Hayes & Dunning, 1997), but when researchers specify a particular definition of a trait, people fail to

⁴We should note that Krueger and Mueller (2002) have disputed this analysis of self-perception among the incompetent, stating instead that the overestimates of poor performers are simply due to a statistical artifact (regression to the mean). Kruger and Dunning (2002) have responded by noting that this concern is minimized, if not eradicated, if one uses measures of performance that are of sufficient statistical reliability.

rate themselves so positively (Dunning et al., 1989), and their ratings begin to agree with those of others (Hayes & Dunning, 1997; Story, 2003).

Information Neglect

People also misjudge themselves relative to others because they ignore crucial information, and this neglect can produce the above-average effect and, on occasion, its direct opposite. It is clear that people often want to know how they stack up against others; in fact, they often prefer to find out how they compare with others than to find out how they stack up against objective standards (e.g., Festinger, 1954; Suls & Wheeler, 2000). Self-views derived from comparisons with others affect people's behavior more than comparisons with objective standards (W.M. Klein, 1997).

Exclusive Focus on the Self With Neglect of Others. Given these findings, it is surprising that people's comparative judgments often involve very little comparison. When evaluating their skill vis-à-vis their peers', people are egocentric, thinking primarily of their own behaviors and attributes and ignoring those of others (Kruger, 1999; Weinstein & Lachendro, 1982). Ask people how well they can ride a bicycle relative to others, and they say they do quite well—mostly dwelling on the fact that they have no trouble riding a bike, but forgetting that other people have no difficulty either. But ask them about their juggling ability, and they describe themselves as worse than average—neglecting again that others are also poor jugglers (Kruger, 1999).

This egocentrism carries important implications. College students, for example, prefer to compete with other college students in a trivia contest focusing on Adam Sandler movies (an easy topic for them) than to compete in one on 19th-century French painting (a hard topic), forgetting of course that what is easy or difficult for themselves would be equally easy or difficult for most competitors. People bet more in poker games as the number of wild cards in the deck increases because they are more likely to have a good-looking hand. But this behavior is irrational because wild cards do not play favorites, and other players are equally advantaged as the number of wild cards expands (Windschitl, Kruger, & Simms, 2003; see also Moore & Kim, 2003).

Controllability and Privacy of Traits. This egocentrism may be most apparent in two different realms. The first is the realm of controllable behavior. People think of themselves as superior to their peers when thinking about traits that are construed as controllable, but not so much when thinking about uncontrollable traits (Alicke, 1985). People consider themselves more cooperative and self-disciplined than others (all controllable qualities), but not necessarily more creative or lively. People believe they are less likely than others to be involved in auto accidents when they are the driver, but not when they are a passenger (McKenna, 1993). In the health domain, people believe they are more likely

than their peers to avoid risks associated with high-fat diets and alcohol (all conceptually under a person's control), but not risks related to pesticides or environmental contamination (Sparks & Shepherd, 1994).

The second realm in which this kind of egocentrism is particularly apparent is that of private traits and behaviors. People tend to believe that they possess traits more than their peers to the extent that those traits tend to be expressed internally. For example, people tend to think they are more self-conscious, self-critical, and choosy than their peers, but not that they are more aggressive, poised, or wordy, traits that are more external in their expression (Miller & McFarland, 1987). They tend to believe they feel emotions more intensely than their peers do (Sabini, Cosmas, Siepmann, & Stein, 1999). They also think they harbor more uncertainties and ambivalences than their contemporaries do. For example, college students believe they are more ambivalent than their peers in their stance toward casual sex (Lambert, Kahn, & Apple, 2003), smoking, and illegal drug use (Hines, Saris, & Throckmorton-Belzer, 2002). People also believe they are more inhibited than their peers (Vorauer & Ratner, 1996). In essence, what people do not seem to know, or at least what they seem to neglect, is the likelihood that the people around them have just as full an internal life of private thoughts, inhibitions, emotions, and uncertainties as they possess themselves.

Note, however, that this neglect does not uniformly lead to a self-aggrandizing above-average bias. To be sure, it often leads people to think they are superior to their peers (e.g., "I am more choosy than others"), but to an equal degree it can lead people to believe they are inferior (e.g., "I am more prone to be embarrassed than others"), as is evident from the examples we have mentioned. The failure to consider that others may share the same feeling as the self results in part from the inherent inaccessibility of others' private states (Miller & McFarland, 1987; for other explanations, see Miller & Nelson, 2002), whether that inaccessibility leads to a favorable or unfavorable self-attribution.

This neglect of the internal lives of others can result in a phenomenon known as *pluralistic ignorance*, in which people believe they uniquely possess a deviant opinion, whether desirable or not, when in reality most people in the community privately share the same opinion. This phenomenon is aptly illustrated by Schanck's (1932) study of a small rural community where virtually everyone publicly condemned alcohol and card playing because of church dogma. Private interviews, however, indicated that the community members did not hold such extreme views, although they mistakenly assumed that other community members' public denunciations of drinking and card games reflected their private sentiments. Thus, each person in the community paradoxically thought he or she was exceptional in thinking that cards and alcohol were not particularly bad. This pluralistic ignorance perpetuated the community status quo because "even if no one believes . . . everyone believes that everyone else believes" (Miller & McFarland, 1991, pp. 287–288).

Explanations for Overly Optimistic Prediction of Events

When predicting future events, people usually follow a natural and simple psychological strategy: consider some future action or outcome (e.g., “Will I lose 20 pounds on this diet?”) and spin possible scenarios that might produce that outcome. To the extent that these scenarios seem plausible, people will be more confident that the outcome will take place (Kahneman & Tversky, 1982). This procedure seems so simple that it might be difficult, at first, to see what complications it might create, but it does create several, once again wrapped around the themes of unknowable and neglected information (for an extended review, see Dunning, in press). The net effect is that people fall prey to two biases. The first is that the predictions they make tend to be too optimistic. The second is that, whether people make predictions that are optimistic or pessimistic, the confidence they place in those predictions tends to be too high, in that their predictions tend not to prove accurate at the rate they expect them to.

Information Deficits

Unknown Situational Details. People often make overly confident predictions, typically optimistic ones but not necessarily so, because they fail to correct for the fact that the details of future situations are often unknown or unpredictable, even though those details may matter. For example, consider those college students who optimistically predicted they would vote in an upcoming presidential election (Epley & Dunning, 2004). Voting depends, of course, on any number of personal qualities—such as the individual’s commitment to civic duty, as well as his or her interest in the campaign. But it also depends on a host of situational details that may be difficult to anticipate. For example, is the student behind on the paper that is due the next day, is the student sick, did the roommate who promised to drive the student to the polling place show up, or did the student’s parents call for a 2-hour chat?

By definition, such situational circumstances are not knowable until election day arrives. Thus, if the student makes a confident and optimistic prediction that he or she will vote and does not take into account that important situational details matter, the student will likely make confident predictions of desirable behavior that later turn out to be unwarranted (Griffin & Ross, 1991). Recent research suggests that people do indeed fail to consider the unpredictability of crucial situational details when they make predictions of future behavior. If anything, they tend to predict as though they can fully anticipate what those details will be. Asking people to fully describe the details of relevant situations, and then assume those details are accurate, does not increase the already high confidence with which they make predictions. However, asking them to describe alternative ways situations may play out, and noting that it is impossible to anticipate how situations will play out in the future, causes participants to hedge their predictions to a significant degree (Griffin, Dunning, & Ross, 1990).

Imperfect Understanding of Emotion, Visceral Drives, and Their Consequences. Situational features are inaccessible in other ways. People often have difficulty predicting how they will respond to situations that have significant emotional or visceral components—leading to predictions of self-behavior that are often too optimistic, but at times too pessimistic. For example, office workers approached just after eating a substantial lunch predict they would prefer a healthy snack, such as an apple, rather than an unhealthy but more filling snack (a candy bar) to be delivered to them at 4 p.m. a week later, even though they know (intellectually) that they tend to be hungry late in the afternoon. However, when the day of delivery arrives, they tend to prefer the unhealthy candy bar over the healthy fruit they predicted they would want (Read & van Leeuwen, 1998).

Similarly, studies of behavior in social situations show that people fail to appreciate the power of fear, anxiety, and embarrassment to shape their future behavior. In one illustrative study, students were asked, hypothetically, whether they would come up to the front of the class to dance to Rick James’s funk classic “Superfreak” if they were paid \$5. Roughly 30% said they would, but when a real opportunity was presented, only 8% volunteered (Van Boven, Loewenstein, & Dunning, in press). Participants more accurately predicted how they would respond to similar opportunities when they were first emotionally aroused, regardless of which specific emotion was involved (Van Boven, Loewenstein, Welch, & Dunning, 2004), suggesting that once put “in touch” with the level of arousal they would feel in an actual situation, people are better able to provide accurate forecasts about how they would behave.

Being asked to dance in front of a class may not be common, but similar losses of courage have been observed in more socially relevant circumstances that entail a good deal of emotion. For example, women often report that they will respond to sexual harassment with anger and assertiveness. But the dominant emotion of people placed in harassing situations turns out to be fear (Woodzicka & LaFrance, 2001), and victims end up responding to harassing behaviors with silence or diversionary talk rather than confrontation (Swim & Hyers, 1999).

People also fail to appreciate how much they act to minimize or mitigate the impact of negative emotional events, believing that the sting of undesirable events will last longer than it does. Assistant professors who are denied tenure believe that the psychic devastation of that event will last far longer than does in reality; voters whose candidate loses the election overestimate how long that outcome will dampen their happiness (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998). People overestimate how anxious being told they are HIV positive will make them and also how relieved being informed that they are HIV negative will leave them (Sieff, Dawes, & Loewenstein, 1999).

In essence, people underestimate how quickly they adapt to emotional events, particularly ones that are negative. People possess emotional “immune systems” that dispel the negative impact of aversive events. They are adept at discounting bad

news, creative in finding the silver lining buried among the clouds, facile at finding distractions that lighten their affect. However, they also seem to be quite unaware of their resourcefulness in protecting themselves from the psychic harm caused by negative events (Gilbert et al., 1998).

Information Neglect

Neglect of Alternative Scenarios. People also mispredict future events because they neglect important information that they have in hand. When they spin scenarios about how they will behave in the future, they tend to dwell on positive scenarios and fail to take into account worst-case scenarios that they could easily generate. The scenarios that people report as the “most realistic” scenarios tend to resemble more the scenarios they report as the “best case” than those they report as the “worst case.” Asking people explicitly to write down a worst-case scenario fails to influence their sanguine predictions about future events, although exposure to another person’s worst case causes them to be more pessimistic about that other person’s prospects (Newby-Clark, Ross, Buehler, Koehler, & Griffin, 2000).

Neglect of Concrete Detail. People also base their predictions about events in the distant future on abstract, higher-level features of a situation and give short shrift to more concrete, low-level features that can have a significant impact on behavior (Trope & Liberman, 2003). Often, whether or not a behavior is desirable is a high-level feature, and concrete details about the feasibility of the behavior are lower level. For example, when college students were asked which of two class assignments they would rather complete 9 weeks later, they tended to prefer the assignment on the topic of romantic love (a desirable, high-level feature), even though the readings for this assignment were written in a foreign language (a concrete, low-level feature). However, just before the due date, students preferred the second assignment, which had an undesirable topic, the attitude concept, but had required readings written in their native language (Liberman & Trope, 1998).

Neglect of Background Circumstances. People also mispredict because their imagined scenarios concentrate too much on the behavior in question and not about seemingly irrelevant swirls and eddies of everyday life that are not conceptually related to the behavior but that may still interfere with their capacity to perform that behavior. In a phrase, people suffer from focalism, basing their predictions on factors that are conceptually related to the behavior in question and acting as though common background circumstances—those swirls and eddies—are irrelevant, even though once people bring to mind those background circumstances, they may recognize how much those circumstances may influence whether they will behave in the manner they predict (Schkade & Kahneman, 1998; Wilson, Wheatley, Meyers, Gilbert, & Axson, 2000).

Consider, for example, the case of people predicting how quickly they will be able to get their holiday shopping done. They are likely to base their prediction on circumstances that are conceptually relevant to the act of shopping (e.g., how many gifts are needed, how many stores must be visited, what traffic will be like on the way to the stores, how many gifts can be bought over the Internet), spinning a scenario for their prediction that seemingly presumes that the act of shopping is somehow hermetically sealed off from the rest of life. However, it is a problem when people concentrate on the focal act of shopping and ignore the knowable fact that background circumstances from everyday life often sneak in to interfere with one’s plans. It is in ignoring these background influences—or at least ignoring the relevance of everyday background circumstances in general—that that prediction falls prey to error.

To be sure, whether or not one gets the shopping done depends on circumstances related to shopping, but it also depends on factors that, on the face of it, are not related to shopping, such as whether children get sick, the weather turns to snow, in-laws come to visit, or friends decide to throw holiday parties. By ignoring these usually knowable background circumstances, people fall prey to the consequences of focalism (Schkade & Kahneman, 1998; Wilson et al., 2000). And if they are not in a position to know which background circumstances will arise, they at least are in a position to know that some unspecified everyday circumstances will prove relevant to their plans, and so should adjust their predictions accordingly.

Consider one extant example from the psychological literature that serves as a simple and direct illustration of focalism. College students were asked to imagine the aftermath of a football game being played by their school’s team and were then specifically requested to predict how much their mood would be influenced in the days following the game if their team won or lost (Wilson et al., 2000). A few days after the game, their actual mood was measured. Participants tended to overestimate how much their mood would be affected by their football team’s fate. Of key note, however, were the predictions of a separate group of participants who first underwent a “defocusing” intervention, before they made their predictions, in which they were asked to consider all the other mundane activities (e.g., eating, going to classes, socializing with friends) in which they were likely to engage in the days following the game. Thereby reminded of the remainder of life lying outside football, these “defocused” participants made predictions about their mood that were less extreme and more accurate.

Neglect of the Lessons of Experience. Perhaps the most relevant information people have in hand is knowledge about their behaviors and outcomes in the past. The intriguing property of the planning fallacy is not that people prove too optimistic about how quickly they will complete projects, but that they prove too optimistic despite a lifetime of experience, to which that they will freely admit, that demonstrates they rarely complete projects

well before the deadline. People ignore this previous experience because they generally take an “inside view” rather than an “outside view” when predicting how quickly they will complete tasks (Buehler et al., 2002; Kahneman & Lovallo, 1993; Kahneman & Tversky, 1979). When people take the inside view, they consider the unique features of the task at hand and imagine a series of steps that will lead them from their starting point to a solution. As a consequence, they focus on their abilities and resources, perhaps envisioning obstacles and thinking about how they will overcome them.

In contrast, when people take an outside view, they dismiss this scenario building focusing on the situation at hand and instead pursue a more data-driven strategy in which they just tally the final outcomes from situations they know of that are similar to the one they now face. Some of these data can come from their own previous experience; some of the data can come from the experiences of other people. For example, to predict whether they would successfully lose 20 pounds on a diet, people taking the outside view would tally the successes and failures they have had with diets in the past, as well as the number of successes and failures their acquaintances have had. They would then make a prediction based on this count. But although the outside view is relevant for making a good decision, the inside view seems much more compelling and natural. “The inside view tells a colorful story; the outside view recites statistics” (Camerer & Lovallo, 1999, p. 315).

Research on the planning fallacy shows that people typically take an inside view of their situations even when outside-view information is available (Buehler et al., 2002). When experimental participants were asked to think aloud about a project, 74% of their thoughts were about the future, and the future they contemplated was a particularly rosy one. Only 3% of the participants spontaneously considered potential problems. Only 7% of them considered information relevant to the outside view of this particular project—their own past experiences with similar projects (Buehler et al., 2002).

No matter how colorless, the outside view contains useful information that might help people make better decisions (Buehler et al., 2002). To predict the future, people are better off recalling the past and assuming its straightforward relevance for the future, rather than focusing exclusively on how the future might unfold. When students in one study were asked to predict when they would complete an academic task, they predicted that they would do it about 4 days in advance of the deadline (a rate that only about 30% achieved). However, when asked when they normally accomplished such tasks, they admitted that they normally finished only 1 day before the deadline—and this held true for the project they were predicting. Similarly, a random sample of Canadian taxpayers thought that in the current year they would mail in their return about a week earlier than usual, but they generally completed their returns about when they did in previous years.

Thus, incorporating the outside perspective has been shown to enhance the accuracy of predictions about the future. For ex-

ample, college students make unbiased predictions about when they will complete an assignment when asked to list their completion times for previous assignments and then reminded that this past experience may prove relevant to the current assignment (Buehler et al., 1994). Lovallo and Kahneman (2003) described a group of academics working on revising the curriculum of a local school system. When they were asked to predict how long it would take the group to finish their job, the most pessimistic prediction was 30 months. One member of the group did concede upon questioning that in his extensive experience it usually took such groups 7 years at best to complete their task, if they completed it at all. The group completed its work 8 years later.

Summary

In sum, a wealth of evidence suggests that people make substantial errors when they evaluate their abilities, attributes, and future behavior. Several psychological mechanisms conspire to produce these faulty self-assessments, but many of them can be sorted into two general classes. First, erroneous self-assessments arise because people often do not have all the information necessary to provide accurate assessments, and they do not take into account what they do not know. Second, erroneous self-assessments arise because people neglect relevant and useful information that they do have in hand.

However, we should hasten to add that our review of the processes underlying flawed self-judgment is not exhaustive. We have highlighted two general themes that recur often in the literature, but there are other classes of psychological mechanisms that are also responsible for errors in self-assessment. To name just one example, people are often motivated to reach flattering conclusions about themselves and their place in the world. Thus, they mold, manage, and massage the feedback the world provides them so that they can construe themselves as lovable and capable people. The psychological literature is replete with demonstrations of all the tricks and techniques people use to construct and maintain desirable images of themselves while avoiding negative ones (for reviews, see Baumeister & Newman, 1994; Dunning, 2001; Kunda, 1990)—and some echoes of these demonstrations are apparent in the material we discuss in the following sections. Thus, it should be understood that there are other processes in play that might lead people to form incorrect impressions of themselves.

In the sections that follow, we turn to real-world domains to see if the patterns and themes we have described are repeated as one focuses more on real-world settings. We place the domains of health, education, and work under close scrutiny to see whether flaws in self-assessment arise not only in the laboratory, but also, for example, in the doctor’s office, the classroom, and the corporate boardroom. In each domain, we review pertinent literature and describe ways in which findings echo, contradict, or inform the basic research findings already described. To presage our conclusions somewhat, we find that the literature in each domain

reaffirms portions of the delineated themes. The literature in each domain also is informative about other forces that produce flawed self-judgment, demonstrates some of the consequences of these flaws, and suggests strategies that might improve the accuracy of self-assessment, or at least alleviate the costs of erroneous ones.

HEALTH

Among the three highly valued resources Ben Franklin referred to in his widely cited advice, healthy precedes wealthy and wise. Public opinion polls show that physical health is consistently rated as one of the most valued of personal resources. The United States spends more of its gross domestic product on health services than any other major industrialized country (Braden et al., 1998).

In light of the value and importance of health, people should be highly motivated to understand their personal risk of developing a serious illness or injury. However, they succumb to numerous misperceptions and erroneous beliefs about personal health risks and the relationship between symptoms and illness. In this section, we first consider biases in perceiving health risk and then discuss how erroneous perceptions of social norms and individual uniqueness affect healthy and unhealthy practices. Then, we discuss how laypeople's potentially mistaken beliefs about health and illness prompt mistaken self-judgments that influence their efforts to obtain health care, as well as their adherence to medical regimens.

Several themes we discussed earlier are echoed in this summary of self-perceptions and health. In particular, people tend to be overly optimistic about their vulnerability to health risks. They also tend to exhibit pluralistic ignorance in ways that affect their health-related behavior. In addition, misjudging the emotion of situations, they tend to mispredict the preferences they would have if they faced certain medical decisions, and thus make decisions for other people that differ from what those other people want. Finally, people at times exhibit confidence in their ability to diagnose themselves, an example of the double curse of incompetence.

Unrealistic Optimism About Health Risks

A wide range of lifestyle behaviors, such as smoking, drinking too much alcohol, overeating, and avoiding physical exercise, are associated with premature death and disease (e.g., Belloc, 1973). To be motivated to relinquish these practices, people need to recognize their personal risk of disease and injury. However, they tend to be unrealistically optimistic about their health—as they are about other areas of life—perceiving themselves to be significantly less at risk than their peers for a wide range of physical diseases and negative health outcomes. For example, when a large sample of nominally healthy adults completed a questionnaire that listed a series of health problems and other hazards (e.g., food poisoning, lung cancer, drug addiction, ulcer, mug-

ging, stroke, serious auto injury, cancer, high blood pressure), they typically rated their personal risk between *average* and *less than average* (Weinstein, 1987).

Studies focused on specific health risks echo these findings. People tend to believe their personal risk of becoming ill from their food (e.g., salmonella poisoning) is lower than the risk for most other people (Sparks & Shepard, 1994). Chinese students and European Canadians in Toronto rated their personal risk of being infected with severe acute respiratory syndrome (SARS) as lower than their peers' risk, with the bias being stronger among the Chinese (Ji, Zhang, Usborne, & Guan, 2004). Almost 90% of gay men not infected with HIV rate themselves as having significantly lower risk of becoming infected than the average gay man (Gold & Aucote, 2003). In the United Kingdom, secondary school students who smoke express unrealistic optimism about their health risks, although they do concede that their risks are higher than nonsmokers' (Sutton & Bolling, 2003).

In sum, there is a pervasive tendency for people to perceive themselves as having significantly lower risk of a host of physical health problems than their peers. Logically, however, everyone cannot have a less-than-average risk. Some individuals, by virtue of their age, occupational level, and education, are at lower risk than others, but Weinstein (1987) found that unrealistic optimism about health risks was independent of these factors. He also found that the health risks people were likely to be unrealistically optimistic about were those associated with the incorrect belief that if the problem has not yet appeared, it is unlikely to occur in the future. This unrealistic optimism may make people feel overly comfortable and reduce their motivation to take preventive action.

Causes of Unrealistic Optimism and Unrealistic Pessimism

Many factors converge to produce unwarranted optimism. Weinstein (1983, 2003) proposed that unrealistic optimism is based on a need to defend self-esteem against possible threats. Several findings point to such a motivational explanation. Typically, if people are asked to generate reasons why a particular outcome might happen to them, or to think through a series of events that could lead to this outcome, their perceptions of the likelihood of this outcome increase (e.g., J. Sherman, Skov, Hertz, & Stock, 1981). This tendency, however, can be biased by people's desire to think of themselves as healthy.

Weinstein and Klein (1995), for example, asked some experimental participants to list all their personal attributes that would tend to increase the likelihood that they would become 30% overweight or would develop a drinking problem. In another condition, different participants listed all their personal attributes that would decrease their likelihood of experiencing these outcomes. In both conditions, participants then estimated the likelihood that they would experience these problems in the future. The pattern of results was consistent with a motivational explanation of unrealistic optimism: Participants who focused on personal factors that should make a problem more likely did not

show the expected reduction in the optimistic bias, presumably because they did not want to think of themselves as being at high risk. However, participants asked to list attributes or behaviors that would reduce the risk of obesity became significantly more optimistic.

Nonmotivational factors may also produce unrealistic optimism—and pessimism. Egocentric neglect of others, for example, comes into play. Because people consider their own chances of experiencing an event but neglect the fact that the average person probably faces the same likelihood (Chambers, Windschitl, & Suls, 2003), they tend to think that they are both more likely to experience common events and less likely to experience rare events than their peers are. This results in unrealistic optimism in the case of common desirable events, as well as in the case of rare undesirable events. However, when people consider rare desirable events and common undesirable events, the tendency reverses itself, and people express undue levels of pessimism (see also Kruger & Burrus, 2004).

Consequences of Unrealistic Optimism

Unrealistic optimism and unrealistic pessimism are problematic for promoting good health, regardless of whether motivational or cognitive processes are responsible for these biases.⁵ One of the core components of many theories about health promotion—from the *health belief model* (Becker, 1974; Rosenstock, 1966) to *protective motivation theory* (Rogers, 1975) to *precaution adoption theory* (Weinstein, 1988)—is that perceived vulnerability is an important motivator for people to do something about their health. Indeed, a meta-analysis has shown that the degree to which people perceive themselves to be vulnerable to health problems predicts the likelihood that they will engage in health-promoting behaviors (Harrison, Mullen, & Green, 1992). Thus, being overly optimistic may dissuade people from learning about or adopting preventive measures because they fail to perceive any personal risk. Alternatively, unrealistic pessimism may prompt feelings of hopelessness that serve as obstacles to preventive actions.

The literature does provide examples showing that perceived invulnerability leads to complacency in the shadow of health risks. When people perceive themselves as relatively invulnerable to flu, they are less likely to intend to obtain a flu shot (Larwood, 1978). People who believe their risk is lower than that of their peers are more likely to engage in high-risk sex (Sheer &

Cline, 1994) and less likely to use contraception (Burger & Burns, 1988) than people who believe they have high risk.

That said, the literature on whether unrealistic optimism uniformly discourages prevention behaviors is mixed (Buunk, Gibbons, & Visser, 2002; Weinstein, 2003). At times, undue optimism is linked to behaviors that might prove problematic; at other times, perceived invulnerability does not affect health-related behaviors. One study even showed that lower levels of optimism were associated with adverse consequences. Rutter et al. (1998) found that among motorcyclists, lower optimism correlated with greater likelihood of abandoning precautions in the future! Perhaps the motorcyclists who perceived their risk more realistically (i.e., less optimistically) assumed that this recognition made them safer.

Moderators of Unrealistic Optimism

Some people, however, are more unrealistically optimistic than others. Experience matters, for instance. Drivers who have been hospitalized after a road accident are not as optimistic as drivers who have not had this experience (McKenna & Albery, 2001). Similarly, middle-aged and older adults are less optimistic about developing medical conditions than their younger counterparts are (Madey & Gomez, 2003), presumably because older persons have had more exposure to health problems and aging. Acutely ill college students (approached at a student health center) perceive themselves to be at greater risk for future health problems than do healthy students, indicating that risk perceptions can be “debiased” if the person has a relevant health problem. Acutely ill students, however, continue to be unrealistically optimistic about problems that do not involve physical health (Kulik & Mahler, 1987).

In a similar vein, within a few days of experiencing the 1989 California earthquake, college students displayed no unrealistic optimism about being hurt in a natural disaster, but they were unduly optimistic about their invulnerability to other kinds of negative events. However, when surveyed 3 months later, these students were unrealistically optimistic about their vulnerability even to natural disasters. This study indicates that although people can be debiased after experiencing negative events, the debiasing seems to be event-specific and short-lived (Burger & Palmer, 1992).

One of the strongest moderators of unrealistic optimism is perceived control (Helweg-Larsen & Shepperd, 2001). The greater a person’s perceived control over an event or its outcome, the stronger the person’s optimistic bias. A meta-analysis conducted by C. Klein and Helweg-Larsen (2002) synthesized the results of 21 studies and found a moderate relationship between perceived control and the bias. For example, people are less optimistic about becoming sick because of air pollution (low control) than about becoming injured in an automobile accident (high control because precautions, such as fastening the seat belt, are possible; McKenna, 1993; Quadrel, Fischhoff, & Davis, 1993).

⁵We should note that we concentrate on the impact of unrealistic self-views for physical health. During the past 15 years, there has been a vigorous debate about the impact of unrealistic self-views on mental health. The dominant assertion in this debate is that overly positive self-views are actually beneficial for coping and psychological adjustment when people face extreme adversity (Taylor & Brown, 1988), such as cancer (Taylor, Lichtman, & Wood, 1984) or the aftermath of a civil war (Bonanno, Field, Kovacevic, & Kaltman, 2002). This assertion, however, is not without its critics (Colvin & Block, 1994; Shedler, Mayman, & Manis, 1993), and everyone agrees that if overrating one’s self is advantageous, it is desirable only in moderate doses (Baumeister, 1990; Taylor & Brown, 1994; see Dunning, 2005, for a review of this debate).

Interventions to Reduce Unrealistic Optimism

Thus, exposure to negative life events can reduce unrealistic optimism about similar events that might take place in the future, at least temporarily. The question is whether physicians and other health professionals might have other techniques at their disposal to make people more realistic about their chances of succumbing to serious illness or injury.

Two techniques have demonstrated some success in reducing unrealistic optimism. One is personalized feedback. Kreuter and Strecher (1995) asked patients waiting in family practice clinics to complete a questionnaire about their perceived and actual risk of dying from heart attack or stroke within the next 10 years. Perceived risk was measured with questions about degree of risk (i.e., respondents indicated whether they thought their risk was much lower than average, lower than average, etc., compared with other people their same age and sex). Actual risk was assessed with questions about age, height, weight, blood pressure, and other risk factors (Amler, Moriarty, & Hutchins, 1989). Algorithms were then used to calculate each patient's actual risk for each cause of death. Two to four weeks after completing the baseline questionnaire, patients received mailed feedback about their actual risk (i.e., whether it was average, lower than average, or higher than average compared with the risk of other people of the same age and sex). Six months later, a follow-up questionnaire on perceived risk (using the same format as in the original questionnaire) was mailed to the patients. Results indicated that patients who were initially unrealistically optimistic about their stroke risk became more realistic after receiving personal risk information. Individualized risk feedback had no effect on perceived risk of heart attack, however. These results suggest that receipt of objective risk information can debias risk perceptions, but more research clearly is needed. This is a promising approach, but probably not appropriate for mass campaigns (Weinstein, 2003). Using personalized risk feedback may be appropriate when there is direct contact with the individual, as, for example, during a medical visit.

The second technique that has been shown to reduce unrealistic optimism about health risks targets the motivational basis of such unrealistic optimism. If one confronts people with information about risk factors, they will defend against it. However, if their sense of self-worth is bolstered before they encounter that information, they become more willing to accept it and to change their behavior. For example, college students who wrote about a personally important value (e.g., how much they cared about their friends and family) before viewing an AIDS-awareness film were more affected by the message than were students in a control group, who wrote about a more neutral topic. They viewed themselves as more at risk and purchased a greater number of condoms as they left the laboratory (D.A.K. Sherman, Nelson, & Steele, 2000). Reminding people of a time in which they were kind to other people, a memory that affirms their self-esteem, makes them more willing to accept information that their health is

at risk, and less likely to distort their memory of that information (Reed & Aspinwall, 1998).

Exceptions to Unrealistic Optimism

The literature, however, also shows that erroneous assessments of personal health risk can sometimes be overly pessimistic. For example, in the study by Ji et al. (2004), although Chinese respondents believed that they were less susceptible to SARS than their peers (i.e., relative risk), they overestimated their absolute risk of being infected. Similarly, a representative sample of Malawians expressed significant unrealistic pessimism regarding their susceptibility to malaria and schistosomiasis (Morrison, Ager, & Willock, 1999), and they also reported feeling they had less control over these diseases than other people do. Interestingly, in settings with endemic health risk, feelings of pessimism and lack of control may actually be adaptive. Indeed, supplementary evidence indicated that persons who perceived the diseases to be low in predictability and controllability were more likely than others to adhere to healthy practices.

In Kreuter and Strecher's (1995) study of personalized feedback, patients were also queried about their risk of cancer and injury in a motor vehicle crash. Respondents were initially unrealistically pessimistic about these risks, perhaps because of all the mass media attention given to breast cancer and traffic accidents. These instances of unrealistic pessimism run against the usual tendency toward undue optimism, and future research could profitably delineate the conditions under which people overestimate their risk rather than underestimate it. For example, one perspective mentioned earlier (Chambers et al., 2003) predicts unrealistic pessimism about common negative events because people fail to appreciate that common events are just as likely to happen to other people as to themselves.

Perceptions of Uniqueness and Consensus

Up to now, we have been discussing how people tend to see themselves as unique among their peers, usually in a self-congratulatory direction. Pluralistic ignorance is a different kind of false-uniqueness error, whereby individuals falsely think their private opinions are at odds with those of everyone else, but do not feel good about being different (Miller & McFarland, 1987; Suls & Wan, 1987). Instead, they feel odd, different, deviant.

Pluralistic ignorance can also play a role in health-related behaviors, such as the excessive consumption of alcohol on college campuses, a major concern of college administrators, health officials, and parents because of the resulting injuries, legal infractions, and low academic performance (Wechsler, Davenport, Dowdall, Moeykens, & Castillo, 1994). Transition to college typically is associated with increases in alcohol consumption, and this shift appears to represent the effects of strong public norms among students favoring excessive rather than moderate consumption (Friend & Koushki, 1984). Students, however, may have several reasons to have personal misgivings

about these practices. They get first-hand exposure to hung-over roommates and inappropriate behavior associated with drinking.

Prentice and Miller (1993) speculated that college students may have real concerns about excessive drinking but feel inhibited about sharing them with peers because nearly everyone's public behavior on campus seems to support excessive drinking as the norm. To evaluate this idea, Prentice and Miller had Princeton University students rate both their own and their friends' comfort level with alcohol practices on campus. Students consistently rated themselves as less comfortable than their friends (see also Perkins & Berkowitz, 1986; Suls & Green, 2003) and also as less lax about the number of drinks they consumed per hour and about the number of drinks they would consume if they were going to drive afterward (Bourgeois & Bowen, 2001).

These misperceptions have behavioral consequences. Believing their concerns are unique, students may set aside their misgivings and strive to conform to the perceived social norm (Schachter, 1951). Indeed, Prentice and Miller (1993, Study 2) found that personal alcohol consumption increased over the course of the semester among students who believed their behavior initially diverged from the perceived norm. College males seem to feel this social pressure more intensely than college females, perhaps because alcohol plays a larger role in their campus social lives and signals a "machismo" image (Suls & Green, 2003). Pluralistic ignorance also has been found with respect to other undesirable health practices, such as marijuana and amphetamine use and unsafe sex practices (Suls & Green, 2003).

Luckily, there has been some success using interventions to correct erroneous perceptions of social norms. For example, freshman college students who were exposed to peer-oriented discussions that focused on pluralistic ignorance relating to alcohol (discussions about the difference between public appearances and private beliefs) drank less than students who received an individually oriented discussion about the risks associated with excessive drinking (Schroeder & Prentice, 1998). Agostinelli, Brown, and Miller (1995) also found that providing realistic feedback about misperceived norms was effective in reducing excessive drinking in college students.

People, however, also make the reverse error, overestimating how much their attitudes, habits, and outcomes coincide with those of other people. Such beliefs can have a powerful influence on behavior (e.g., Ajzen & Fishbein, 1980). In essence, the difficulty of identifying true social norms may force people to depend on simple rules of thumb. A heuristic that is often used to judge the frequency of events is based on availability (i.e., in memory)—an event or behavior is perceived to be more common the more memorable or accessible to memory it is (Kahneman & Tversky, 1973). Because people's own behavior and attributes are most accessible to them, they may distort norms in the direction of their own behaviors or preferences. This phenomenon, the false-consensus effect, has been well documented in a wide range of domains (Mullen et al., 1985; Ross, Green, & House, 1977).

For example, Chassin, Presson, Sherman, Corty, and Olshavsky (1984) found that adolescents' ratings of their friends' use of alcohol, cigarettes, and marijuana were positively correlated with their own personal use. Suls, Wan, and Sanders (1988) reported a false-consensus effect for a wide range of health-relevant behaviors (e.g., substance use, use of seat belts). For example, college students who smoked cigarettes believed that more of their peers smoked cigarettes than did nonsmokers. In contrast, nonsmokers perceived there to be more nonsmoking students than smokers did.

These erroneous perceptions have behavioral effects. For example, among smokers, overestimating the prevalence of smoking is directly associated with an increase in this behavior (Gibbons & Gerrard, 1995). Not only does the false-consensus effect justify one's personal practice, but the erroneous perception may perpetuate and even increase the practice. Luckily, false-consensus effects can be corrected with special health-promotion programs. For example, Hansen and Graham (1991) developed a school intervention that involved providing students in junior high school with true information about the prevalence of drinking, as well as perceptions of acceptability. This intervention reduced alcohol, marijuana, and tobacco use among the students over a year's time. Other researchers have had success with similar programs in secondary schools and colleges (Borsari & Carey, 2001; Perkins, 2002).

Using the Self to Predict the Preferences of Other People

False consensus can also influence the decisions family members make when a spouse or other relative is too sick to make medical decisions for him- or herself (Fagerlin, Ditto, Danks, Houts, & Smucker, 2001). In such cases, the family member is asked to use the standard of *substituted judgment*, that is, to make the decision that the incapacitated patient would make if he or she could. The idea is that the surrogate should choose only those treatments the patient wants. But if people tend to believe erroneously that others share their behaviors and opinions, a surrogate's decisions about life-sustaining treatments may reflect the surrogate's preferences more than the patient's.

More to the point, if surrogates are not able to accurately project what their own preferences would be if they were in the patient's place, then they will make decisions that contradict the patient's actual preferences. The potential for this type of misjudgment is not trivial, given that people tend to misperceive how they would respond to emotionally laden situations when they are not currently emotionally aroused (Van Boven et al., 2004, in press).

Fagerlin et al. (2001) most directly showed that surrogates make false-consensus errors when predicting the preferences of patients. The researchers read six end-of-life scenarios to senior citizens and to the persons they would designate as their surrogates in joint sessions. As the older adults made treatment decisions for each scenario, they were asked to describe their

rationale for making each decision, and the surrogates were asked to raise any concerns or questions they had. Following the discussion, both parties individually completed a questionnaire involving life-support preferences and predictions in several realistic illness scenarios (e.g., Alzheimer's disease with no chance of recovery, terminal colon cancer with no pain, stroke with a slight chance of recovery).

The results indicated that surrogates' predictions of the senior citizens' preferences more closely matched their own treatment wishes than the wishes of the senior citizens. Supplementary analyses showed that the patients' preferences also contributed to the surrogates' predictions—indicating that the patients had input into the decision-making process—but that the surrogates' own preferences also significantly influenced the decisions. The assumption of similar treatment wishes might be a conscious one on the part of surrogates, but also may occur with little awareness. The most important lesson to be taken from this research is that the standard of substituted judgment may be impossible to fully meet because of the influence of the false-consensus effect and people's inability to project accurately what their own wishes and desires would be if they were in the patient's position.

Symptom Perception and Commonsense Models of Illness

It would seem clear that medical diagnosis should be left to the professionals. If a person is not sure about what to make of his or her chest pain or indigestion, it would make sense for that person to seek out a doctor's expert judgment about what might be wrong. This logic is simple and direct, but it leaves out factors that influence the decision about whether a trip to the physician is warranted. Herein lies the problem: People hold an arsenal of commonsense theories about health and illness that they endorse with confidence, and these theories determine whether they seek out medical care or adhere to what their doctor recommends. These theories, however, may be quite wrong—and given the research on self-insight into one's incompetence (e.g., Kruger & Dunning, 1999), there is reason to believe that people feel quite competent in their role of playing doctor even when this self-confidence is misplaced.

People's labeling of and responses to symptoms are strongly determined by the commonsense models of illness that they hold. These models are based on personal experiences regarding symptom onset, duration, and intensity (Leventhal, Meyer, & Nerenz, 1980). People also observe patterns of covariation between symptoms and illnesses and use these patterns to attribute symptoms to possible causes. These beliefs are supplemented by information and experiences provided by family members and friends.

As an illustration of how symptom perception leads to decisions about whether to seek care, consider a study by Johnson and King (1995), in which they interviewed heart attack survivors. Results indicated that people experiencing symptoms of a heart attack took longer to seek medical intervention if the symptoms

did not match their beliefs or expectations about what a heart attack "should" feel like. Similarly, chronically elevated blood pressure (i.e., hypertension) is asymptomatic—people have no signals they can check to tell whether or not their blood pressure is elevated. However, a majority of patients (from 46% to 94%) diagnosed with hypertension believe that they can tell when their blood pressure is up (Baumann & Leventhal, 1985; Meyer, Leventhal, & Gutmann, 1985). Some patients think their pressure is high only when they are headachy; others think it is high when they are feverish, and so on. These patients fail to recognize that hypertension is a chronic condition.

These misperceptions may also influence compliance to prescribed medications. For example, hypertensives who believe their pressure is high only when they experience headaches may take their medications at only those times. Meyer et al. (1985) found that hypertension patients were more likely to adhere to their physician's recommendations (e.g., medication, exercise, change in diet) when they believed their medication was effective in controlling unpleasant symptoms than when they felt their symptoms persisted, even though these symptoms had no veridical relation to blood pressure. Interestingly, the patients acknowledged that "most people" cannot detect variations in blood pressure; they thought they personally were exceptions to the rule. Good compliers also tended to recognize that hypertension is a chronic or cyclic disorder. In contrast, patients who dropped out of treatment believed they had an acute disease that had been "cured" by temporary changes in diet and medication.

People follow a number of rules of thumb in their commonsense models of illness. These rules may seem compelling, but they often are in error. One of these heuristics is the *symmetry rule*; people expect symptoms to denote illness, and they expect illness to be accompanied by symptoms. Clearly this belief is often true, but overuse of this rule increases the probability of diagnostic errors. It can lead people to believe that they are well if they are not experiencing symptoms. This rule is misapplied in the case of common and serious diseases that exist without symptoms, such as hypertension. Another consequence of the symmetry rule is that people who believe they are ill expect to perceive symptoms. In a study consistent with this idea, Baumann, Cameron, Zimmerman, and Leventhal (1989) gave participants in the laboratory bogus feedback indicating they had elevated blood pressure readings. Subsequently, the participants complained of headaches and flushing—symptoms they erroneously believed were associated with hypertension.

Another heuristic is the *prevalence rule*, according to which health threats that are rare or unusual are also serious, and commonplace symptoms are mild or harmless (Jemmott, Croyle, & Ditto, 1986). The consequences of the prevalence rule are nicely documented in a medical anthropology study of Hispanic migrant workers who ignored serious gastrointestinal symptoms. Interviews indicated that the workers concluded that the symptoms were harmless because they were so common in their community (Clark, 1959).

The *stress-illness rule* (Cameron, Leventhal, & Leventhal, 1995) is a common heuristic applied to symptoms and illness (Kelley, 1967). According to this heuristic, if symptoms covary with troublesome life events, the symptoms may be attributed to stress rather than illness. People are particularly likely to attribute symptoms to stress if they are vague or ambiguous, rather than dramatic (e.g., profuse bleeding). For example, students who are preparing for an examination may discount physical symptoms because they assume that the symptoms are due to the exam stress.

Very recent research has provided some initial evidence that these erroneous theories about disease can be corrected. One such study of heart attack patients found commonsense models about heart disease could be changed through customized educational efforts—and that patients receiving such education were subsequently less incapacitated by chest pain, left the hospital sooner, and even walked on their own sooner than patients who did not receive this intervention (Petrie, Cameron, Ellis, Buick, & Weinman, 2002).

Mistaken Commonsense Models of Illness Based on Stereotyping

Mistaken notions of disease can come from two other everyday sources. One is media attention. In a recent review about symptom perception, R. Martin and Leventhal (2004) noted that the *New York Times* (Brody, 1999) reported an abrupt increase in visits to gastroenterologists by people with chronic heartburn shortly after a highly publicized study in the *New England Journal of Medicine* reported that chronic heartburn was associated with an increased risk of developing esophageal cancer (Lagergren, Bergstrom, Lindgren, & Nyren, 1999). This increase in physician visits did not arise because there had been a sudden increase in heartburn symptoms among the population. Rather, mass media attention made the heartburn symptoms more salient to people and provided them with a new label that represented a potentially serious health threat. Prior to this media blitz, these patients simply labeled their heartburn as a reaction to spicy foods.

Symptom ambiguity makes people highly susceptible to such situational and social influences. For example, if someone mentions to subjects that it is flu season, they are more likely to report relevant physiological symptoms, presumably because being primed with the idea of “flu season” leads subjects to attend to ambiguous physical signs (such as scratchy throat and headache; Pennebaker, 1982). Another example is medical students’ disease (Mechanic, 1972), whereby as medical students learn about the symptoms of various diseases, many of them come to believe incorrectly that they have contracted one or more of them at one time or another. This may result from health-related symptoms being primed and made temporally accessible as a result of reading about medical disorders.

A reliance on social stereotypes also can lead people to erroneous commonsense notions of illness. Consider heart attack symptoms (e.g., chest pain, shortness of breath, sweating, and shoulder pain). When people experience these symptoms, they should promptly seek medical attention because early medical intervention such as thrombolytic therapy (clot-busting treatment) is essential in preserving viable cardiac function. However, almost half of heart attack victims delay for longer than 2 hours before seeking medical attention, and a large minority delay for longer than 4 hours. Several large-scale studies have found that women delay longer than men after the onset of cardiac-related symptoms (Dracup et al., 1997). This sex difference in treatment delay may explain why outcomes tend to be worse for women than for men; extended delay reduces women’s eligibility for thrombolytic treatment. The extended delay for women, however, is curious in that women tend to be more responsive to most health threats and report more physical symptoms than men (Pennebaker, 1982).

These differential behaviors, and outcomes, for women versus men arise because commonsense models of illness also include stereotypes, which in the case of heart disease may lead to an individual using gender as a guide to whether medical attention is needed (R. Martin, Gordon, & Lounsbury, 1998). Specifically, symptoms such as chest pain, shortness of breath, and sweating are more likely to be attributed to a possible heart attack when the victim is a male rather than a female. Cardiac disease is more common, prior to age 65, among men than women. Therefore it is easier to think of a man with coronary heart disease than a woman with the same (Friedman & Rosenman, 1959). If women and men have a commonsense model of heart disease as being a male disease, then women who experience cardiac-like symptoms are apt to discount them as indications of stress or some other disorder. In contrast, men who experience the same symptoms will likely assume that they are cardiac in origin.

To test this hypothesis, R. Martin et al. (1998) presented participants with information about a person reporting chest pain and other symptoms consistent with an acute heart attack. Information about this target person (gender, high vs. low life stress) was manipulated, and participants were asked to rate the likelihood that the symptoms were due to a cardiac illness. For male targets, the symptoms were interpreted as representing heart disease, regardless of whether he recently experienced high- or low-stress life events. However, participants responded very differently to information about the female target, even though her symptoms were identical to those reported by her male counterpart. In the low-stress condition, the female target’s symptoms tended to be attributed to heart disease. But when the female target had high stress, participants were less likely to attribute her symptoms to cardiac causes. This selective discounting was found in college students, adults, and physicians and was not affected by how participants indicated their opinion (i.e., by using rating scales or writing narrative, open-ended responses to the vignettes; R. Martin & Lemos, 2002). Additional

evidence for the gender-based stereotype is that recall of biographical details regarding a male (stereotype-consistent) victim is superior to recall of such details regarding a female heart attack victim (R. Martin et al, 1998, Study 4).

An important question is whether gender stereotypes about cardiac-related symptoms influence whether women experiencing the symptoms of an evolving heart attack seek treatment. R. Martin et al. (2004) explored this issue in a sample of heart attack survivors. Although the men and women patients were of comparable age and risk status and had similar medical histories and symptoms, the women were less likely than the men to attribute their prehospitalization symptoms to cardiac causes. Also, among the patients who received opinions and advice about their symptoms from friends and relatives, women were less likely than men to be told that their symptoms might indicate a heart attack and were also less likely to be advised to seek medical attention. In an attempt to correct these tendencies, the American Heart Association currently sponsors the “Go Red for Women” campaign, designed to raise consciousness about the prevalence and early signs of heart disease in women.

EDUCATION

Similar issues in self-assessment arise in the classroom. Students, obviously, profit from self-assessment that is accurate. Whether it be a junior high school student wondering whether to crack open his social studies textbook one more time before the test or a medical student deciding whether to practice her intubating technique before the next shift, students make more effective decisions about where to apply their learning efforts when they can accurately discern their strengths and weaknesses (Thiede, 1999; Thiede, Anderson, & Theriault, 2003). Accurate self-assessments allow students to become more autonomous agents in their education, taking responsibility for gaining and improving on their knowledge and skill (Boud, 1995; Dochy, Segers, & Slujsmans, 1999; Topping, 1998).

Accurate self-assessment is valuable all the way up the educational ladder. However, it is especially crucial in higher education and professional school settings, particularly as some schools move to a problem-based or case-based model of instruction (Barrows & Tamblyn, 1980; Boud & Feletti, 1991). This approach to education has gained numerous adherents in medical schools, and its prevalence has recently increased in many other settings in higher education (e.g., Duch, Gron, & Allen, 2001). In problem-based learning, students working in groups learn by confronting real-world cases and are responsible for figuring out how to approach each case to bring it to a reasonable resolution. This approach places a premium on problem-solving skills, professionalism, and learning in hurly-burly circumstances that mimic real life.

An essential component of problem-based learning is that students must identify what skills they need to acquire and what knowledge they must gain—in short, they must make correct self-

assessments of strengths and deficits (Boud, 1995). Accurate self-assessment is also crucial for education to be a lifelong enterprise that continues far after the student has left the classroom (Guest, Regehr, & Tiberius, 2001; Sambell & McDowell, 1997). Company executives must continue to educate themselves in a changing business environment; doctors must refresh and enhance their treatment skills; and airline pilots must continue to monitor their flying proficiency.

Against this background, it is sobering to see that although self-assessment of skill and knowledge does tend to bear some relationship to objective performance, the relationship tends to be meager to modest. In a review on academic performance, Hansford and Hattie (1982) discovered that the average correlation between academic self-views and actual performance lay around .21. In a more extensive review, Falchikov and Boud (1989) found that, on average, the grades that college students would give to their work correlated .39 with the marks their teachers would give. Students also showed signs of overinflated self-views, in that 68% of the time they gave themselves higher marks than their teachers would.

To be sure, some specific circumstances did lead to improved, albeit still imperfect, self-assessments. Self-assigned grades, for example, were slightly more related to teachers' evaluations when the solutions to course assignments became more well-defined (e.g., in science classes vs. other areas of study). In addition, the correlations between grades that students gave themselves and teachers' grades were higher in advanced classes than in introductory courses. Studies containing more rigor in their design revealed more student-teacher agreement than those of lesser quality (Falchikov & Boud, 1989).⁶

Three observations should be made about the fact that advanced students provided more calibrated self-assessments than introductory students. First, this increase in calibration was not simply due to students' age. Older students were no more calibrated than younger ones. Second, this increase was observed for students who were assessing their classroom performance, and there is suggestive evidence that as students leave the classroom to face more unsettled and demanding training settings, their self-assessments do not necessarily become more accurate. For example, Arnold, Willoughby, and Caulkins (1985) tracked self-assessments of medical students from their first year, spent primarily in the classroom, to their final year, spent mainly in clinical settings. The relationship between students' self-ratings and supervisors' evaluations fell as students progressed through their studies. In addition, during the final year, students' self-ratings failed to correlate at all with their board scores.

⁶An anonymous reviewer of this manuscript made the insightful point that student-generated grades were occasionally used as inputs into formal grades in some of the studies examining student-teacher agreement. To date, the impact of this circumstance on student-teacher agreement is unknown, but it would be interesting to examine whether attaching consequences to student-generated grades prompts flaws or accuracy in the evaluations students produce.

Third, although experience prompts students to provide more accurate self-assessments, these assessments are still far from perfect—and not all students profit from experience, even when that experience is repeated and provides clear-cut feedback. Hacker, Bol, Horgan, and Rakow (2000) tracked students as they took repeated exams in a course, asking them before each test how well they thought they would perform. Better students provided more accurate predictions as the semester wore on, but poorly performing students did not, remaining dramatically overconfident despite the feedback they had received on previous exams.

The modest accuracy of self-assessments of academic performance is also striking given data on peer assessment, which indicate that assessments by peers are more accurate than self-assessments. In an extensive meta-analysis, Falchikov and Goldfinch (2000) found that an evaluation from a single fellow student correlated .72 with the teacher's evaluation, and that grades given by peers on average tended to be no higher than those given by teachers. To be sure, one should compare these results and those for self-assessment with caution, in that this meta-analysis on peer assessment involved different studies and circumstances than the one on self-assessment (Falchikov & Boud, 1989); students may have taken assessments of their peers more seriously than they did self-assessments.

However, in studies that specifically focus on self- and peer assessments in the same class, peer assessments tend to correlate more highly with instructors' evaluations and objective performance measures than do self-assessments (e.g., Lennon, 1995; Sullivan, Hitchcock, & Dunnington, 1999). For example, among surgical residents, Risucci et al. (1989) found that peer assessments correlated more highly with supervisors' evaluations and performance on an objective test of surgical skills than did self-assessments. Indeed, self-assessments did not correlate whatsoever with performance on the objective test.

These results suggest that when it comes to self-assessments among students, there is a good deal of room for improvement. The issue of self-assessment in education is complex because imperfections in self-assessments are produced by many different problems, depending on the circumstances. However, recent evidence from the literatures on cognitive and educational psychology suggests two general themes that explain erroneous self-judgments. One theme has to do with common educational practices that may confer the appearance but not the reality of skill. The other has to do with people's ability to judge whether they comprehend what they have just read. Both the cognitive and the educational psychology literatures also provide hints of procedures that students and instructors might follow to improve self-assessment.

Educational Methods That Undermine Accurate Self-Assessment

The goal of education is to impart learning—to give students knowledge and skill that they remember and can call upon at

some later date in circumstances that may differ from those of the classroom. Learning, thus, possesses two components. One is *retention*, the ability to recall information or perform a skill over the long term. The second is *transfer*, the ability to apply the knowledge or perform the skill across a number of relevant situations. Whether the topic is American history, playing the cello, or interpreting the nuances of tax law, instructors train students with the goal that those students will be able to retain and transfer the lessons learned to some unknown but relevant circumstance at some date in the future.

With this goal in mind, one common instructional method provides a paradox. The method that most effectively promotes the rapid acquisition of knowledge and the highest levels of proficiency at the end of the lesson—and thus the appearance of learning—is the one that ensures that whatever is learned in the classroom will be forgotten rapidly. The method of instruction that produces these effects is *massed training*, in which instructors train students in one or a few intense sessions. Massed training has advantages. Students undergoing intense training quickly obtain the relevant skill and then display it at a high level. Indeed, much research shows that massed training is more efficient than any alternative for bringing students quickly to a high level of performance (Dempster, 1990; Glenberg, 1979, 1992). In addition, other practices can be combined with massed training to further promote rapid knowledge gain and high performance. For example, students learn more quickly if instructors keep the conditions of learning constant and provide continuous feedback (Bjork, 1999), and if instructors model a solution or performance for students before they begin to generate their own (such as when a tennis instructor models a correct backhand stroke; Bjork, 1988; Jacoby, 1978).

There is only one central problem with this recipe. Although massed training produces quick learning and high performance in the short term, the knowledge and skill imparted tend to be forgotten rapidly—a result that has been known since the time of Ebbinghaus (1885/1964). In essence, although massed training is efficient for acquiring skills, it is not optimal for retaining them. Nowhere might this problem be more evident than in training people how to drive cars. Although millions of dollars are spent on formal driver-education courses, research data suggest that such courses do not produce safer drivers (Mayhew & Simpson, 1996). Indeed, evidence tenuously suggests that young people who complete formal driver-education courses experience more frequent accidents and injury than those who learn more informally from their friends and family (Skelly, 1968; Stock, Weaver, Ray, Brink, & Sadoff, 1983). Similarly, in Norway, truck drivers who were formally trained in skid control on slippery surfaces experienced more crashes, not fewer, than those not trained (Christensen & Glad, 1996), presumably because formal training did more to raise drivers' confidence than their competence (Wilde, 1998).

Retaining knowledge and skill requires a different recipe. Students retain information and skill better when they acquire

information and practice their skill via *spaced*, or *distributed*, training, in sessions that are divided over several occasions, even though this means that students initially learn more slowly and with more difficulty (Dempster, 1990; Glenberg, 1992). Retention is enhanced even further if instructors reverse other practices that are associated with rapid acquisition. Students retain information more successfully if instructors “change up” the circumstances under which material is learned, adding some variability and unpredictability to the presentation of material (e.g., Reder, Charney, & Morgan, 1986). For example, baseball batters learn to hit curveballs, fastballs, and sliders better if those pitches are thrown to them in a random sequence rather than in blocks (e.g., first fastballs, then curveballs, then sliders) during practice (Hall, Domingues, & Cavazos, 1994). Students also retain information better if instructors at times withhold feedback and avoid providing a model that students can imitate (Bjork, 1999).

But if distributed training is superior to massed training when it comes to retention, why does massed training continue to be popular? Part of the answer lies in a common illusion shared by instructors and students. People confuse speed and ease of learning with competence—and this confusion leads to errors in self-assessment. Students and instructors both assume that if a skill has been learned quickly and the student finds it easy to perform, then the student will maintain the skill in the long term (Bjork, 1994, 1999). Short-term excellence is mistaken for long-term competence.

Two examples show how massed training misleads self-assessment. First, Baddeley and Longman (1978) taught postal workers in Britain how to type so that they could use a new mail-sorting system. Workers learned to use the system under one of four different training schedules that ranged in how massed versus distributed they were. In the most massed schedule, employees learned to type in two 2-hour sessions per day for 20 days. In the most distributed schedule, employees trained for one 1-hour session per day for 60 days. Trainees were more satisfied with massed training than they were with distributed training, even though the most distributed schedule produced the best long-term retention of typing skill.

Second, D.A. Simon and Bjork (2001) asked college students to learn keystroke patterns either in a blocked format or in a random one that interweaved the patterns. Participants learned the keystroke patterns more efficiently with blocked presentation than with random presentation—and as a consequence were more optimistic about what their performance would be later on. However, when participants were tested the next day, participants trained on the random schedule outperformed those who had learned under a massed format.

Problems in Reading Comprehension

Much of learning involves picking up a textbook and reading it over, taking care to commit to memory important facts and principles presented in the text. Obviously, good study skills

require knowing when one has comprehended the material in the text and can remember it later when tested, but research presents a cautionary tale about the ability of students to know when they have understood text material adequately and committed it to memory.

People are not very good at assessing their comprehension of written materials. They think they have understood a piece of text when they have not. Glenberg, Wilkinson, and Epstein (1982; see also Epstein, Glenberg, & Bradley, 1984) asked college students to read several passages and to rate their understanding of each. Various sentences in the passages directly contradicted one another, yet students failed to recognize this fact—leading them to express high but unwarranted confidence in their comprehension of the materials.

Other studies affirm this pattern. In a common experimental paradigm, students are asked to read a number of short texts and to predict how well they could answer questions about each. These predictions are then compared with actual performance on quizzes probing comprehension of and memory for the material. Researchers then correlate each participant's predictions and performance. In 1998, Maki identified 25 studies that had used this paradigm; the average correlation between prediction and performance was .27—statistically significant, but hardly strong. Furthermore, poor readers and lackluster students have an especially difficult time knowing when they have understood a text (Bol & Hacker, 2001; Maki & Berry, 1984; Maki, Jonas, & Kallod, 1994; Moreland, Miller, & Laucka, 1981; Shaughnessy, 1979; Sinkavich, 1995; although see Glenberg & Epstein, 1987, for an exception).

Blind spots in assessing comprehension create problems because students regulate their study habits on the basis of what they think they have (or have not) understood and committed to memory (for a review, see Son & Metcalfe, 2000). If students think they have successfully absorbed what is in chapter 3 of the textbook, they will put it aside to stare down chapter 4 (the one that confused them last time they read it). Students who know what they understand study more effectively than those who misjudge their comprehension, and they perform better on exams (Dufresne & Kobasigawa, 1988, 1989; Thiede, 1999). For example, Dufresne and Kobasigawa (1989) asked children in first, third, fifth, and seventh grade to study booklets filled with paired-associate tasks (i.e., the students had to memorize a target word that was paired with each cue word), and they were supposed to study the booklets until they could remember all the word pairs perfectly. Some of the booklets were designed to be easy (i.e., the words in each pair were related to one another), and some were difficult (e.g., the words were not at all related). Fifth and seventh graders studied the difficult booklets more than the easy ones, but their younger counterparts studied the two kinds of booklets equally. When tested later, older children outperformed their younger peers—more often achieving perfect scores—exactly because they had spent their study time more prudently on the harder booklets.

Improving Self-Assessment—and Thus Enhancing Learning

In sum, extant research suggests that students face two general problems when it comes to assessing their own skills. One is that massed training may mislead them into thinking they have acquired a skill that they will retain. The other is that it is difficult for them to assess accurately whether they have understood and can remember what they have read. More recent research, however, has begun to suggest some practices that instructors or students can use to improve self-assessment, as well as actual academic achievement.

Introducing “Desirable Difficulties” to Instruction

Massed training, along with associated practices, produces rapid learning that is difficult to retain, leaving students potentially overconfident in their knowledge and ability. One simple way to alleviate this problem is to introduce “desirable difficulties” that harm the speed with which students learn but that leave them better able to retain what they have learned and to transfer it to different situations in the future. Bjork (1994, 1999) discussed what “obstacles” to place in front of students to enhance long-term proficiency. These obstacles include spreading training over several sessions, varying the circumstances of the training, reducing feedback, and providing “contextual interference” (i.e., practicing different subskills under a random schedule rather than under a blocked one).

Forgoing massed training and introducing such desirable difficulties, however, might be somewhat difficult for an educational institution to do (Bjork, 1994, 1999). The rapid improvement and high short-term performance that massed training provides look impressive to instructors and their administrators. When formally evaluated against alternatives, massed training wins out on many desirable criteria—performance is high, time is used effectively, and students express high levels of satisfaction. Instructors and administrators fail to recognize the long-term shortcomings of massed training because instructional programs based on this method tend not to be evaluated over the long term. Providing long-term evaluations of retention and transfer would go a long way toward identifying when massed training provides more the illusion of competence than the fact (Bjork, 1994, 1999).

Long-term evaluation of retention would also go a long way toward identifying potentially important instances when enduring learning is not possible and therefore skill training should not be attempted. For example, some scholars have proposed that some forms of driver’s education might do more harm than good, making drivers more confident about their ability to handle difficult driving situations than is warranted. As a substitute, they have proposed that drivers forgo skills training and instead undergo “overconfidence training,” in which students are shown just how difficult it is to handle adverse driving conditions so that they drive more carefully—or just stay home when it is wise to do so (Wilde, 1998).

Improving Reading Comprehension via Self-Testing

Students can also take several steps on their own to improve their self-assessment of their reading. One is to test their understanding as they read text, rating what they believe their understanding to be before they test that understanding. For example, Koch (2001) asked physics students to stop periodically what they were reading to rate how well they thought they had understood material they had just read. The students then answered a few questions about what they had read and noted any discrepancies between their ratings and actual performance. Students who went through this exercise posted higher scores on a subsequent physics test than students who did not engage in this form of self-questioning (see also Walczyk & Hall, 1989, for similar results). However, self-testing can sometimes be hazardous to one’s academic health if the self-tests do not accurately represent the subsequent performance situations. Bol and Hacker (2001) had students take a practice test before taking an actual exam a few days later. Relative to a control group, students who had taken the practice test performed worse, presumably because the practice test was not representative of the material covered on the actual exam.

Self-testing is particularly valuable if students and instructors take care to place a delay between study and the self-testing exercise. Self-testing done immediately after study tends to lead to inaccurate self-assessments. Why might this be so? Students tend to make assumptions about their memory on the basis of the accessibility or fluency of material (Benjamin, Bjork, & Schwartz, 1998) even when that fluency is not diagnostic of long-term memory. Just after studying, material is quite accessible and fluent, and people are not in a good position to predict what material they might forget. However, after a delay, students need not lean on their intuitions about memory. Instead, they have actual data about the information they have retained and what they have forgotten.

Several studies show that self-assessments are more accurate if collected after a delay rather than right after study. For example, Thiede and Dunlosky (1994) asked college students to remember the translations of several Swahili words. The students were presented with the words again either immediately after studying them or after approximately a 30-second delay and asked whether they would recall the translations later in the experimental session. Self-assessments provided after the 30-second delay more accurately predicted subsequent recall than did self-assessments collected immediately (see Dunlosky & Nelson, 1992; Kelemen, 2000; Nelson & Dunlosky, 1991), a result that has been found to be just as true of kindergartners as it is of college students (Schneider, Vise, Lockl, & Nelson, 2000). Indeed, the more the delay the better (Kelemen & Weaver, 1997).

The fact that judgments of learning are more accurate after a delay should inform other interventions for improving the accuracy of self-assessment. Thiede and Anderson (2003) discovered that asking students to summarize a passage they had recently read led to more accurate self-assessments of learning—but only

if students delayed a short while before summarizing the text. Similarly, Thiede, Anderson, and Therriault (2003) found that simply asking students to write down five keywords from material they had studied improved self-assessments of learning, but again only if there was a delay between study and the keyword exercise. Moreover, among students allowed to study the material a second time before the self-assessment, those who had undergone the delayed keyword exercise outperformed their peers who did not do the exercise or had done it right after their initial study session.

Reviewing Past Performance

Reviewing one's performance has been shown to lead to better self-assessment of a variety of skills. For example, medical students reach a more appropriate appreciation of their interviewing skills if they watch videotapes of their performance (Ward et al., 2003). This videotape exercise can be enhanced when students review the videotapes with faculty (Lane & Gottlieb, 2004; Scherer, Chang, Meredith, & Battistella, 2003). The review of past work need not necessarily involve videotape. Cochrane and Spears (1980) had students taking a clinical dietitian course periodically rate their own performance and then meet with a faculty member who had also rated their performance. Over time, students' self-assessments came to correspond more closely with those of their supervisors.

Benchmarking

Self-assessments may also be improved when people take pains to compare their choices and performance against those of others, a practice known as benchmarking. For example, Farh and Dobbins (1989) asked college students to play the role of editor, correcting an error-strewn report that someone else had putatively written. After completing the task, students rated their performance. Some students, however, were shown other students' efforts before being asked to provide self-assessments. The self-ratings of these students correlated more strongly with objective markers of performance than did the self-ratings of the students who had no benchmarking opportunity.

Benchmarking has also been found to improve the self-ratings of medical students. In one study, family-practice residents completed a standard exercise in which they interviewed a mother who might have physically abused her child and subsequently rated how well they thought they had performed along several dimensions. They then watched their own videotaped interview, along with four benchmark interviews that displayed a wide range of competence in interviewing technique (D. Martin, Regehr, Hodges, & McNaughton, 1998). Correlations between self-ratings of performance and supervisors' ratings were higher after students had viewed these benchmark interviews than before.

An important caveat regarding benchmarking is that it may improve the self-ratings of high-performing students, permitting them to see just how special their skills are, but not the self-assessments of poor-performing students. Kruger and Dunning

(1999, Study 3) suggested this when they asked students with excellent grammar skills and those with poor grammar skills to review grammar tests that had been filled out by other students. After this benchmarking exercise, top performers appropriately raised their estimates of how distinctive their grammar skills were relative to those of other students, but poor performers did not revise their self-estimates in any meaningful way. Echoing this result, Hodges et al. (2001) reanalyzed previous data on medical residents and discovered that benchmarking exercises helped top performers to recognize the excellence of their interviewing skills—in that their self-ratings after benchmarking more closely corresponded to their instructors' ratings—but did not prompt poor performers to recognize deficits in their interviewing skills. In short, benchmarking might help students in general gain insight into their skill, but it does not help the poor performers who most need to adjust their self-assessments.

Peer Assessment

Recent work also suggests that students benefit from peer assessment, presumably because such feedback informs students of shortcomings they do not necessarily know they have. Academic feedback from peers can take on many forms, from informal criticism to formal grades. Studies of peer assessments have examined a wide variety of academic tasks, although work in this area is still in its infancy (Topping, 1998). Researchers have looked at peer assessments for writing, oral presentations, group projects, and professional skills. The research results suggest that peer assessments are generally reliable and correlate highly with teachers' evaluations (Falchikov & Goldfinch, 2000; Topping 1998). Students whose work is evaluated by their peers tend to achieve higher academic marks than those who do not go through peer-assessment exercises. Students also tend to possess more positive attitudes toward the learning experience after peer assessments exercises (Topping, 1998).

The success of peer assessment in raising achievement should perhaps not come as a surprise. Peer assessment might improve students' accomplishment via many different mechanisms. Beyond improving students' imperfect evaluations of their own work, peer assessment creates more "time on task," allows students to practice their skills, engages them more fully in their assignments, prompts them to reflect more on what they have done, and assigns them more responsibility and accountability (Topping, 1998). Important side benefits may include enhanced social skills, in that students learn to give and receive feedback (Marcoulides & Simkin, 1991).

Peer assessment may prove especially valuable in cases in which structured and formal education is neither preferred nor even available. Such a case arises in schools that use a problem-based learning model, which leaves students more on their own during their studies. Perhaps a more important case arises when people leave formal schooling, yet still must engage in continuing or lifelong learning. In both these circumstances, students must decide for themselves whether they need further education

(Topping, 1998). In the medical realm, peer assessment, coupled with self-assessment exercises, may prove a significant help with health care training in developing countries where formal supervision is expensive to maintain. In the late 1990s, the Indonesian government, in concert with several international organizations, conducted a training program with local health-clinic counselors to improve their communication skills when discussing contraception with clients. Counselors attended a 5-day refresher seminar and then were assigned to one of three groups. For the next 16 weeks, one group completed self-assessment exercises on their communication skills. The second group completed self-assessment exercises, plus attended weekly peer-review sessions. The third group served as a control and did not receive any intervention.

Six months after training, researchers taped and coded two interviews involving each counselor, to assess the extent to which each counselor provided medical information and exhibited a facilitative and active communication style. Both interventions increased facilitative communication—with the intervention involving peer assessment producing more improvement in how much information (e.g., about lifestyle consequences) clients were given. As a consequence, clients talked more actively with counselors who had received the intervention that combined self- and peer assessment (Kim, Putjuk, Kols, & Basuki, 2000). The costs associated with these self- and peer-assessment interventions turned out to be low given the effects these interventions brought about.

Peer assessment, however, is not without difficulties. It can be demanding and time-intensive. It can provoke anxiety (Topping, 1998), as well as fears that grades will be more a product of illegitimate influences (such as friendship) than of objective attainment (Dancer & Dancer, 1992). Poor students do not provide as accurate assessments as their more skilled peers do (Saavedra & Kwun, 1993). Thus, using peer assessment is not guaranteed to improve student performance unless the assessments are conducted carefully. Peer assessments become more valid as they are based on a larger number of observations and a greater number of dimensions of skill. They are also most helpful when standards are clear and more than one peer provides an assessment. Peer-assessment exercises are also enhanced if instructors communicate the purpose of the exercises clearly, articulate the dimensions of judgment clearly, provide training when necessary, and monitor students' evaluations, intervening when they are too harsh or too lenient (Norcini, 2003). Peer assessments better match teachers' evaluations when students provide global evaluations based on a number of well-defined criteria rather than a number of separate evaluations along several individual dimensions (Falchikov & Goldfinch, 2000).

THE WORKPLACE

Given the research reviewed earlier in this monograph, it is not surprising that biased self-views also affect success and social

relationships in the workplace. When one looks at the accuracy of self-assessment in the workplace, from the office cubicle to the executive boardroom, one sees that people tend to hold overly inflated self-views that are modestly related to actual performance. One also sees reverberations of other themes (overconfidence in judgment, egocentric neglect) that are apparent in the domains of health and education.

If people often have biased self-views, it is not surprising that these biases affect workplace outcomes ranging from the ability to anticipate performance evaluations to the ability to gauge promotion opportunities. What is perhaps less apparent is that organizations can adopt many routines and procedures that keep biased self-views from affecting performance in many situations. As a result, biased self-views may have their most systematic and damaging effects in the circumstances in which individuals are least likely to face routines and procedures that could repair their biased decisions, for example, when people confront new projects that have few precedents, or when CEOs and management teams at the very tops of organizations, where there are few routine correctives, make important, large-stakes decisions about which new markets to enter and which companies to acquire.

Accuracy of Self-Knowledge in Organizations

In organizational life, the largest surprises generated by lack of self-knowledge may be those that are produced when self-evaluations are not echoed by supervisors who set raises and hand out promotions. If employees overrate their own performance, it is difficult to imagine how people could not wind up disappointed at least some of the time. For example, Zenger (1992) studied several hundred engineers at two high-tech companies and found that 32% of the engineers in one company and 42% in the other rated their own performance in the top 5% of all engineers. Imagine the difficulty of conducting honest performance evaluations for these engineers.

Although workers may find it tempting to blame their supervisors when they receive a less-than-stellar evaluation, research suggests that they should trust their supervisors' views more than their own. Typically, the views of other people—subordinates, peers, and superiors—agree with each other more often than with self-views. In one typical study, Bass and Yammarino (1991) studied U.S. Navy officers who had graduated from the U.S. Naval Academy and were on active duty on ships ranging from tenders to aircraft carriers. The researchers collected leadership ratings from the officers' subordinates and the officers themselves. They also retrieved some measures of leadership from the officers' superiors—each officer's performance on regular "fitness reports" and promotions relative to peers. On all seven positive leadership dimensions studied (e.g., charisma, individual consideration, intellectual stimulation), the officers rated themselves more positively than did their subordinates. But although the officers' self-ratings displayed almost no correlation with

their own fitness reports and promotions, the subordinates' ratings inevitably did.

This is one example, but its message is echoed across many studies in many kinds of organizations. Harris and Schaubroeck (1988) conducted a meta-analysis of 30 years of academic research comparing the self-ratings of individual employees with the ratings of their supervisors or peers. Although the ratings of peers and supervisors agreed pretty well ($r = .62$), their correlation with people's own ratings of their job performance was lower ($r = .35$ for supervisors' ratings, $r = .36$ for peers' ratings). This lack of agreement between self-ratings and outside ratings was exacerbated when there was more ambiguity: when the rating scale was ill-defined (e.g., correlations were lower when people assigned a single global rating rather than separate ratings on specific task dimensions, and they were lower when people rated abstract traits rather than specific behaviors) or when job descriptions were ambiguous (correlations were lower for less-well-defined managerial or professional jobs than for blue-collar or service jobs).

For a number of reasons, individuals may find it difficult to learn about how well they are performing (Ashford, 1989). First, the environment may be tricky. Organizational standards of performance may be shifting, complex, or subject to disagreement. Second, individuals may avoid seeking feedback—getting feedback requires them to balance the potential benefits of improvement against the cost of learning something that would be a blow to self-esteem. Also, individuals may resist seeking feedback even if they might like to have it because asking for feedback might make them appear insecure or needy.

Challenges in Providing Employees With Feedback

If people are not very good at assessing their own performance, and if they are reluctant to solicit feedback on their own, cannot organizations adopt routines (e.g., regular performance reviews or evaluations) that could clue employees in to their strengths and weaknesses? Although the organizational literature remains perpetually optimistic about solving this problem by encouraging regular and reliable feedback, actual feedback systems tend to be relatively ineffective for the following reasons:

- Feedback is *infrequent*: A study by the American Management Association in 1984 found that 75% of organizations provided formal feedback only once a year (Ashford, 1989). When feedback is this infrequent, it takes on additional freight and significance, exacerbating the other problems in this list.
- Feedback is *threatening*: Because people are overconfident about their abilities, true feedback will often be perceived as negative and, potentially, threatening. Feedback may not work if people experience it as commentary about their character or future outcomes (e.g., "Is she saying I've been a bad manager?" "Will I lose my bonus this year?") rather than about the way they are approaching a specific task (e.g., "Let's think of ways that your team could produce 40 more widgets a week"; see DeNisi & Kluger, 2000; Kluger & DeNisi, 1996).
- Feedback is *sugarcoated*: People generally find it painful to receive negative feedback, so evaluators find it painful to give negative feedback. Thus, evaluators may withhold negative feedback or distort it in a positive direction.
- Feedback is *given too late*: Because evaluators are hesitant to disillusion the people they need to evaluate, they may wait so long to say anything that by the time they do, they are beside themselves with anger (Larson, 1989).

Various solutions for making feedback more effective have been suggested and explored. Some researchers have argued that feedback should be separated from evaluation, and that feedback should be provided more frequently (e.g., "coaching"); but in a busy organization where formal reviews are done at most once a year, asking for less formal feedback multiple times a year might be unrealistic. Other researchers have advocated taking the burden of reviews off one individual manager (e.g., by using 360° reviews, in which a focal individual may be evaluated by peers, subordinates, customers, and multiple superiors), but this makes the evaluation process more costly and intense. Some organizations have tried to force managers to rate their employees along a curve—only a small percentage of people can be rated as high performers, and half of employees must be rated below average—but such systems frequently produce frustration and controversy (Baron & Kreps, 1999, chap. 10).

Also, because of the basic problems of self-assessment, any of these proposed solutions is subject to its own potential problems. Evaluators may give feedback infrequently because they find it emotionally taxing to give people feedback that may violate their self-image. The burden of feedback systems falls especially hard on front-line managers, who have to give feedback that disappoints or angers employees who assume they are above average and cannot see why their managers do not agree. Giving people feedback more frequently, although it may reduce the stakes and allow for more specific instruction that would be less threatening, forces evaluators to repeat a painful process more often. Because of these problems with giving feedback, many organizations evolve to an equilibrium pattern of feedback that may be sub-optimal for organizational performance, but more viable from the standpoint of interpersonal relationships.

Organizations must select the best people for further development and promotion, but as we have indicated in our discussion of evaluation and feedback, this selection process is likely to be contentious when a substantial majority of workers think they are above average. How, then, can organizations recognize and develop their best employees without undermining the motivation of the bulk of employees in the middle? One way might be to give high and low evaluations to a few exceptionally good and bad performers, respectively, and to give the bulk of the employees ambiguously positive evaluations. Figure 1 shows an actual distribution of evaluations for a high-technology company studied by Zenger (1992). Note that relatively few people (around 10%) were given the very top rating, but even fewer (less than 5%)

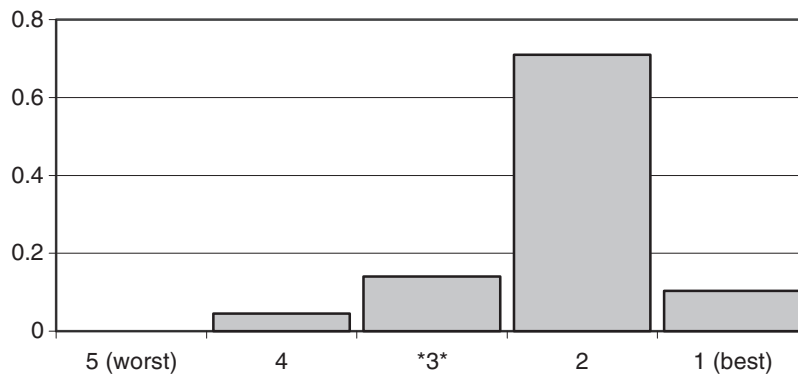


Fig. 1. Performance evaluations at a high-technology company (Zenger, 1992). The graph shows the percentage of employees given each rating on a scale from 1 (top performance category) to 5 (bottom category).

were given ratings (4 or 5) that were explicitly below average. The majority of employees were clustered in a large category that, at least in terms of the overt scale, rated them solidly “above average.” Here, as in Garrison Keillor’s Lake Wobegon, everyone can be above average; but this organization gives people evaluations that prove it.

Zenger (1992) noted that such rating systems may reduce employees’ concerns with the evaluation system. When any particular employee is singled out for special positive recognition, this makes many other employees—all those who correctly or incorrectly feel they are performing better—feel that their efforts have not been equitably rewarded. Organizations with positively skewed evaluation systems may get away with singling out a few really high performers—those who peers agree are clearly deserving—if they avoid tougher distinctions among people in the middle of the distribution, where quality distinctions may be more contested. Such evaluation systems cleverly allow organizations to select the best (and get rid of the worst) with the fewest ruffled feathers, but the problem is that these systems do not provide people in the middle of the distribution with an evaluation that may signal them to improve their performance. The overall results of such systems may depend on whether organizations can separate feedback from evaluation and provide feedback that encourages people in the middle to improve even though they are already being evaluated as “above average.”

Cognitive Repairs for Mistaken Self-Judgment

For people at the bottom of a hierarchical organization, lack of self-knowledge may hinder career or personal success. This is a recipe for personal unhappiness, but it may not affect how they complete their day-to-day tasks. Luckily, organizations often develop or evolve reminders, routines, and procedures that help mitigate problems that could be caused by employees who lack awareness of their own abilities: Marketing managers have their ad campaigns checked by higher-level managers who have created campaigns before, and loan officers in banks must have their loans approved by loan audit committees. Combined, these

standardized organizational routines and cross-checks may prevent employees from making big mistakes through their own lack of self-knowledge.

Adding Safety Factors and Buffer Time

For example, engineers are often overconfident about their ability to calculate the amount of concrete needed for a dam, or the strength of materials needed for an airplane wing. This lack of self-knowledge could have obvious negative implications for people’s safety, but the engineering profession has evolved a “repair” for this overconfidence: safety factors (Heath, Larrick, & Klayman, 1998). Thus, after engineers calculate the amount of concrete they need to hold back a given amount of water for a dam, they multiply this number by a safety factor, a number between 3 and 8, thereby tripling, or even octupling, the amount of concrete their calculations suggest. Are engineers overconfident? Yes. Are they overconfident by a factor of 8? No.

Similar repairs have evolved in other areas. The need for repairs of overconfident estimates in, say, software development is evident DeMarco (1982), a respected consultant on software development, defined an estimated completion time as “the most optimistic prediction that has a non-zero probability of coming true” (p. 14). Software developers at Microsoft often experience burnout because they “grossly underestimate” how long it will take them to accomplish particular tasks (Cusumano & Selby, 1995, p. 94). This hurts the programmers who struggle to accomplish overly difficult goals, and it hurts the organization that finds itself unable to deliver the products it promised at the time it promised them. Microsoft repairs this problem by adding buffer time to projects. In planning the timeline for developing stand-alone applications such as Excel and Word, Microsoft builds in buffer time equal to 30% of the schedule. For more complex projects, such as operating systems, it might add buffer time of 50%.

Microsoft has also developed processes that help individual programmers assess their own self-knowledge more correctly. Said one manager quoted by Cusumano and Selby (1995), “The classic example is you ask a developer how long it will take him to

do something and he'll say a month, because a month equals an infinite amount of time. And you say, 'Okay, a month has 22 working days in it. What are the 22 things you're going to do during those 22 days?' And the guy will say, 'Oh, well, maybe it will take two months.' Even by breaking it down into 22 tasks he realizes, 'Oh, it's a lot harder than I thought' " (p. 254). Liberman and Trope (1998) showed that when people consider tasks that are distant in time, they analyze them at a higher level of abstraction than they do tasks that are close in time; the repair for this problem at Microsoft—breaking an overall task down into 22 subtasks—is a textbook application of Trope and Liberman's (2003) research, but it developed as an organizational practice long before researchers had recognized and elucidated the psychological mechanism.

Forcing People to Pay Attention to the Environment

Other repairs help people pay attention to the situational forces that affect their outcomes. People habitually take credit for successes, whether caused by their own actions or by situational luck (Mullen & Riordan, 1988), and so do traders on Wall Street who are making particularly high-stakes decisions. Traders have evolved a simple repair to force each other to pay attention to the broader market. They tell each other, "Don't confuse brains and a bull market" (Odean, 1998).

Providing Feedback From Other People

Often, overly optimistic self-assessments can be repaired by forcing optimistic individuals to answer questions from other people who are not as personally committed to a project and are therefore better able to analyze it objectively. The norms of science emphasize rigorous, thoughtful attention to detail, but even well-socialized scientists cannot always recognize the limits of their research. One study that compared successful and unsuccessful microbiology labs found that successful labs, where many scientific breakthroughs occurred, placed more emphasis on group lab meetings in which researchers were forced to answer the questions of their interested but skeptical peers from diverse areas of research (Dunbar, 1995). Disney holds the "Gong Show" for new ideas, during which (as in the 1970s game show) participants bang a gong at a point where a pitch is becoming unconvincing, and the Pentagon requires that important missions be vetted by the "Murder Board," a group of senior combat officers who will kill a mission that is not adequately planned (Heath et al., 1998).

The Problem of the New

The fact that many organizations have evolved cognitive repairs to correct the distorted self-views of their employees suggests that inadequate self-knowledge creates problems for organizations, and that those problems are important enough for some organizations to have invested the effort to design useful correctives. But not all problems are equally likely to be repaired, and current

research does not indicate how often problems are effectively repaired. On the basis of research on individual and organizational learning, we speculate that repairs are most likely to develop when tasks are repeated, when feedback is relatively prompt, and when other employees are in a position to alert the focal decision maker to his or her own lack of awareness (Heath et al., 1998). Interestingly, this suggests that problems with self-knowledge are likely to be particularly pronounced when organizations undertake novel tasks and when top managers, who do not face strong social feedback from peers, are the ones taking action.

Novel Projects

Organizations are unlikely to be able to correct for the inadequate self-knowledge of their employees when they engage in a task for the first time. Existing repairs are unlikely to be effective for truly novel tasks, and the organizational literature is filled with anecdotes about novel, large-scale projects that provide evidence of the planning fallacy writ large. For example, the Sydney Opera House embodied a novel design that stretched the architectural and engineering sophistication of its day. In 1957, the building was predicted to open in 1963 at a cost of \$7 million. After the project was scaled down, the opera house eventually opened a decade later, in 1973, at a cost of \$102 million (Buehler et al., 2002).

But such planning disasters are not limited to single isolated cases. When researchers have systematically examined arenas in which organizations undertake novel projects, they have found that such projects frequently go very wrong. One study of pioneer process plants (i.e., the first plants built to produce chemicals using a new chemical process) showed that the typical plant experienced actual construction costs that were double the original estimate. A study of start-ups showed that more than 80% fell short of their projected market share (Davis, 1985). The U.S. General Accounting Office estimated that when the military purchased equipment involving new technology, it was delivered on time and within budget only 1% of the time (Buehler et al., 2002).

Such broad-scale planning disasters could conceivably be created by a number of social, institutional, and economic factors, but research demonstrates that lack of self-insight, and the planning fallacy specifically, can cause such surprises even when completing a novel project is largely under the control of one individual. One cultural influence that potentially exacerbates the planning fallacy, particularly in the case of novel projects, is socially shared views of management. According to a classic definition, managers "plan, organize, direct, and control" (see Mintzberg, 1975). This definition focuses heavily on the inside view; it encourages managers to picture themselves in charge, directing and controlling what happens. But an important part of self-insight is realizing what one does not control. One good antidote to excessive optimism is realizing, "It's not about you." When people focus on themselves and what they intend to

do, they typically do not consider the large number of ways the environment can react to hinder the best-laid plans.

Not only may definitions of management exacerbate problems with self-knowledge, but the frameworks people use to assess situations may do the same. For example, the field of business strategy provides advice to top managers about how to make major decisions about which new markets to enter. One type of strategic analysis, industry analysis, focuses on the characteristics of various industries that may make them profitable or not (e.g., the power of suppliers or customers, threats from current competitors or potential ones; e.g., Porter, 1980), whereas other types focus more on the internal skills and capacities of the firm (e.g., the resource-based view, discussed by Barney, 1991, or core competencies, discussed by Prahalad & Hamel, 1990). Although both types of strategic analysis are important, psychological research suggests that analysis focusing on the characteristics of industries is likely to be a more effective repair to the initial assessment tendencies of top managers. Left alone, they are more likely to take an inside view, considering the core competencies or resources of their firm, than they are to take an outside view and consider the situational dynamics and history of their industry.

Entering New Markets

If lack of self-knowledge is most damaging when tasks are novel and when there are few social corrections for mistakes, then, ironically, the major negative impact of limited self-knowledge will happen at the upper levels of organizations. Organizations that have elaborate procedures and committees to avoid making a bad \$10,000 loan may have no routines or social correctives to prevent their CEOs from making a bad \$100,000,000 decision to acquire another firm or enter a previously untested market.

Most new businesses fail quickly—61.5% exit within 5 years, and almost 80% within 10 years. Why would entrepreneurs enter markets when their businesses have such a small chance of success? One possible reason is that good business opportunities do not last very long. Businesses that “fail” may actually be rational gambles that take advantage of a short-lived market opportunity. Or perhaps entrepreneurs enjoy taking risks; they may be willing to gamble on a small probability of a fabulous success. A third possibility is that lack of self-awareness plays an important role in decisions to enter new markets: Perhaps entrepreneurs are aware that the average firm is unprofitable, but think that they have unique skills that will allow them to succeed even if most other people fail.

In a series of clever experiments, Camerer and Lovallo (1999) provided evidence that supports this last possibility. They used laboratory-based economic games that participants played for substantial sums of money. The games forced participants to make decisions about entering competitive economic markets; a small number of those who entered a market would earn profits based on their competitive “rank,” but everyone else would lose money. In one market with 8 participants, for example, the

highest-ranked competitor would win \$33, the next-highest would win \$17, and everyone else would lose \$10.

The key experimental manipulation was whether competitive rank was determined randomly (by drawing a number) or through skill (at, say, solving logic problems or answering trivia questions). On the basis of previous research on lack of self-knowledge, Camerer and Lovallo (1999) predicted that people would enter the markets too often when competitive rank depended on skill. Overall, in economic markets where rank was determined randomly, the typical participant was profitable 77% of the time, but when rank was determined by skill—and participants could tell themselves why their own abilities might make them better than others—the typical participant was profitable only 40% of the time.

This effect was magnified when participants decided to participate in the experiment after being told up front that the outcomes in the experiment would be based on skill. Some participants signed up for “an economics experiment” without learning anything else about it; others signed up for an experiment in which their earnings would be based on their skill at answering sports or current-events trivia, and they were told that “people who are very good might earn a considerable sum of money.” Participants in the latter experiments were reminded during their experiment that all the other people in the room had also nominated themselves on the basis of their own expertise, but these self-selected participants performed much worse than the participants who simply signed up for an economics experiment: Non-self-selected participants were profitable in 71% of the skill rounds, but the self-selected participants ended up profitable in only 6% of the skill rounds, thus losing money 94% of the time.

These experimental results serve to explain some behavioral patterns among real businesses. Although most new businesses fail, the failures are not evenly distributed. Moore and Cain (2004) noted that some industries—such as retail clothing stores, restaurants, and bars—are marked by persistently high rates of entry and exit. Although people typically rate themselves above average, this tendency is reduced or reversed in domains that are commonly regarded as difficult (e.g., juggling or chess; Kruger, 1999). As we explained earlier, people tend to focus myopically on their own abilities when making comparisons, so they think they will succeed in easy domains and fail in difficult domains, regardless of the skills and abilities of their competitors.

Moore and Cain (2004) argued that in industries believed to be easy, many entrepreneurs will think they have what it takes to succeed. They studied this hypothesis in a laboratory game, similar to the market-entry game just described, in which ranking was based on performance answering difficult questions (e.g., “What is the closest star outside our solar system?” Answer: Proxima Centauri), performance answering easy questions (e.g., “What is the common name for the star inside our solar system?” Answer: the Sun), or a randomly generated number. On average, too many people entered the economic markets when ranking was based on answers to simple questions, and as a result the typical

participant lost money. However, when ranking was determined by answers to difficult questions or by random drawing, fewer people entered the markets, and the typical participant made money. Interestingly, the losses on the simple rounds occurred even though people predicted that many others would also enter—they just thought they would do better than others at the simple task. This suggests that opportunities for success may be more common in difficult domains than in easy domains. Easy domains attract too many people, whereas hard domains scare people away. In both cases, people would do better if they thought seriously about their potential competitors and not just their own skills and abilities.

Landier and Thesmar (2003) studied the real-world impact of self-assessments in a sample of 23,000 French entrepreneurs. They measured overconfidence by comparing how entrepreneurs forecasted the growth of their firms with the firms' actual results several years later. If the entrepreneurs overestimated the future growth of their firms, they were classified as overconfident. Entrepreneurs were more likely to be overconfident when they were working to implement a new idea rather than a standard business venture (knowledge of the situation is likely to be worse in the case of a new idea) and when they expressed a greater desire for personal autonomy (a result that perhaps reflects a motivational bias).

Landier and Thesmar (2003) found that overconfidence had some positive effects. For example, during the initial stages of their businesses, entrepreneurs who were overconfident worked harder in looking for potential customers and in researching technology, relative to those who were not so confident. But overconfidence also had disadvantages. Entrepreneurs who were overconfident were also more likely to finance their projects with the money they raised informally from themselves, friends, and family as opposed to a bank loan, but when they got a bank loan, they were more likely to rely on short-term (as opposed to long-term) borrowing. Both of these financing options gave the entrepreneurs more control in the short term, but they also exposed the entrepreneurs to greater long-term risk. When problems occurred, overconfident entrepreneurs did not abandon projects quickly enough, and many ended up losing money. In contrast, other entrepreneurs, especially those who had borrowed money from banks, were more likely to terminate their project in time to avoid losing their own money.

In general, psychologists have assumed that optimism and overconfidence are fairly general phenomena that most people share. But by ignoring variation in optimism or overconfidence, psychologists miss the fact that when situations allow people to decide whether to participate or not, this self-selection may ensure that the people who decide to participate are the most optimistic. When someone is among the most optimistic of many contenders, this person's optimism need not be a personality variable, relatively constant across situations; it may simply indicate that the individual has made a mistake in analyzing a specific situation. Interestingly, among Landier and Thesmar's

(2003) French entrepreneurs, education increased overconfidence. In general, one might think that education would provide knowledge that would reduce overconfidence, but because educated people have good options other than starting a new business, the only highly educated people who will give up their existing, already-attractive jobs will be those who are most confident.

Self-selection is particularly important in understanding how common overconfidence will be in social situations in general. Researchers have shown that optimistic self-assessments can be reversed with tasks that are sufficiently difficult. So predicting whether people will be more often optimistic or pessimistic comes down to a question about the environment—is the world filled with tasks that are perceived as easy (which prompt optimism) or tasks that are perceived as difficult (which prompt pessimism)? If environments typically allow self-selection, optimism is probably more likely than pessimism. Even if the world is filled with tasks that are, in reality, hard, people are likely to pursue only the ones they think look easy. Many people may not believe they can sing opera or win programming contests, but many may think they could run a pretty good restaurant or bar. In the laboratory, psychologists can randomly assign people to tasks, even hard ones, but in the world, people self-select situations that seem to fit their self-assessed skills and abilities. Self-selection becomes particularly important when people choose to compete with others who also think they are quite capable. Egocentric neglect becomes more important when people face competitors in a more highly selected group. Senior managers have survived many rounds of a "promotion tournament," and CEOs have survived them all.

The Overconfident CEO

CEOs illustrate the problems of overconfidence in the starkest form. They are highly self-selected, so they are particularly prone to overconfidence. And because they sit at the top of their firms, they have fewer checks on their decisions and face fewer organizational repairs, so any overconfident decisions they make may go uncorrected. In this section, we consider the role of overconfidence in CEOs' decisions about how to invest free cash flows and when to acquire other companies.

Investing Free Cash Flow

In one study, Malmendier and Tate (in press) measured overconfidence by looking for CEOs who were heavily invested in the stock of their companies and who resisted opportunities to sell. Economic theories would predict that such CEOs make the wisest decisions for their companies because their own money is heavily tied up in their firms. But Malmendier and Tate argued that these CEOs are also implicitly expressing high confidence—perhaps overconfidence—in their own ability to manage their firms.

Research in corporate finance has identified an interesting bind faced by CEOs who are overconfident about their own

abilities. On the one hand, they feel great confidence in their ability to start projects that will create economic value, but on the other hand, they find it difficult to raise the money to start those new projects. Financial markets may not share their overly optimistic opinions of their prospects and abilities, and may make it difficult for them to raise money. The stock market may place a low price on their stock, making it hard for them to sell more stock to raise money. Or banks may impose a high interest rate on bank loans to compensate investors for the risk that overly optimistic projects might go into default and not make enough to repay the loans. Thus, overconfident CEOs see a world filled with promising opportunities, but they find it frustratingly difficult to raise money to pursue these opportunities because the financial markets do not share their optimism.

One way CEOs can escape this bind is by starting new projects using only the cash that their organizations generate internally (Heaton, 2002; Malmendier & Tate, in press). Some firms generate more *free cash flow* (i.e., cash that is not committed to paying off loans, etc.) than others, and overconfident CEOs who run such firms have more cash available to devote to “promising” projects.

Malmendier and Tate (in press) found that CEOs who are heavily invested in the stock of their companies are the ones most sensitive to free cash flows—they invest more in new projects when internal cash flows are high and less when internal cash flows are low. By using internal cash to finance their projects, these CEOs avoid having to get outside commitments from financial markets that may not agree with their optimistic views. Interestingly, CEOs are also more likely to invest free cash flows when they hold two other titles, president and chairman of the board. When CEOs are allowed to accumulate all these titles, this may be an indicator that they face weak oversight by their boards of directors. In sum, overconfidence may cause CEOs to see projects as promising even though other people think those projects are risky. And they are more likely to pursue those projects when they can do so without a second opinion from the external financial markets or an independent, active board of directors.

Mergers and Acquisitions

Acquisitions, in which the management of one firm pays to take control of another firm, are a particularly interesting test of self-knowledge (Malmendier & Tate, 2003). Implicitly, the management of the acquiring firm is claiming that they can manage the acquired firm *better* than its current management does. Moreover, they often pay a substantial premium to take over the other firm, and this premium serves as a proxy for how much better they think they can do. Between 1976 and 1990, such takeover premiums averaged 41%, and many were over 100% (Hayward & Hambrick, 1997). These figures are a substantial statement of self-confidence, roughly equivalent to the managers of the acquiring firms claiming that they can manage the acquired firms 41% better than the existing managers. This level of confidence is rarely echoed by external observers. Indeed, on the day most

mergers are announced, the combined stock price of the two firms involved generally *falls*, indicating that the market—composed of people who are voting with their money—predicts that on balance the combined firm will be less healthy than the two firms would be separately (Hayward & Hambrick, 1997; Malmendier & Tate, 2003).

Hayward and Hambrick (1997) investigated the idea that corporate acquisitions might be driven by managerial hubris (Roll, 1986), using a sample of acquisitions that were each valued at over \$100 million. They measured three variables as proxies for hubris. The first was recent organizational success. Hayward and Hambrick argued that such success might foster hubris because people often take credit for their success even when it is caused by luck or general situational conditions. Leaders, in particular, get disproportionate credit for successes (Meindl, Ehrlich, & Dukerich, 1985). The second proxy was media praise, which might accentuate hubris because “romantic” media portrayals might foster the CEO’s self-impression that he or she is clever and in control. Media portrayals might also crystallize and cement the CEO’s power inside the firm as other people come to believe in his or her abilities. The third proxy was relative power within the firm, as measured by the ratio of the CEO’s compensation to that of the next most highly paid corporate officer. CEOs typically receive 30 to 50% more than their closest colleagues, but are often paid more than twice as much. A relatively high ratio means that the CEO’s hubris is unlikely to be checked by social interaction with “inferiors.”

Hayward and Hambrick (1997) discovered that all three measures of hubris independently predicted increased premiums for acquisitions. For example, for every positive article written about a CEO in the *Wall Street Journal* or *Business Week*, the acquisition premium increased, on average, 4.8% (this translates to a large sum of money given that the acquisitions in this sample started at \$100 million). Hayward and Hambrick also showed that the premiums CEOs paid were not justified. The CEOs who received media praise, after all, might have been better CEOs than the managers they replaced; but on average, immediate returns to the acquisitions were –4%, and 1-year returns were –11%. Moreover, CEOs who paid higher premiums generally turned in worse performance as measured by immediate and 1-year returns.

Hayward and Hambrick (1997) treated their measures as indications of some underlying character trait, perhaps even pathology. A psychologist would likely explain the measures more as describing situations. Research suggests that most people are overconfident much of the time (e.g., Dunning et al., 1990; Fischhoff et al., 1977), but social situations often help to mitigate this potential hubris. Loved ones gently talk people out of crazy ideas, and peers give frank feedback when people overstep their abilities.

Thus, it is tragic but not surprising that CEOs, whose power means they have fewer social correctives for their own lack of self-knowledge, may be particularly prone to the perils of hubris.

They are most likely to make bad decisions when their social situations conspire to give them even more power—when a series of successes makes it appear that they are on a roll, when their next-closest colleagues have much less power (as illustrated by lower paychecks), or when the external media grant them more status by endorsing their past decisions.

In other research on acquisitions, Malmendier and Tate (2003) showed another effect that may be produced by a lack of self-insight. CEOs who held a large number of stock options in their own companies were more likely than other CEOs to acquire firms whose major lines of business differed significantly from their own. Economists classify mergers into two categories depending on the similarity of the firms that are being merged. *Related* mergers join two companies in the same general business (e.g., two electronics firms), whereas *diversifying* mergers join companies in different businesses (e.g., a railroad company and a steel producer). Compared with other CEOs, CEOs who are heavily invested in the stock of their own companies are more likely to conduct diversifying mergers—which indicates that they think their management skills extrapolate well to very different industries. The stock market, filled with independent observers who vote with their money, takes an especially dim view of such mergers—stock prices drop on average when any merger is announced, but diversifying mergers produce a reaction that is especially negative.

Interestingly, the problems of self-knowledge illustrated by overly optimistic CEOs may be apparent to other people (at least outside the CEOs' own organizations). Malmendier and Tate (2003) also found increased acquisitiveness by CEOs who were described by the financial press as "confident" and "optimistic," as opposed to "reliable," "cautious," or "conservative." These "confident" CEOs were 78% more likely to conduct a diversifying merger than less confident CEOs.

Mitigating Overconfidence

From our viewpoint, what is interesting about the problems of self-knowledge illustrated by Malmendier and Tate's (2003) CEOs is that the standard economic "solution" to the problem does not solve the problem and even may even make it worse. Traditionally, economists have worried that CEOs might engage in "empire building," taking on too many projects or acquiring firms when they should not, because they are seeking to selfishly acquire additional status and prestige from managing a bigger firm. The standard solution to this problem is to give CEOs a bigger financial stake in their firms. Presumably, when CEOs own a significant amount of stock, they have a stronger incentive to make decisions that benefit all the firm's shareholders because they themselves are major shareholders.

But Malmendier and Tate (2003) found their largest effects on free cash flow, acquisitions, and diversifying acquisitions with CEOs who were heavily invested in their own firms. Our psychological analysis suggests that the problem is not one of incentives, but of self-knowledge. Presumably, Malmendier and

Tate's CEOs did not *intend* to waste their own money by undertaking diversifications, but thought they were doing something good for themselves and other shareholders. These results are consistent with psychological research on decision making, which indicates that adding incentives to a situation does not solve a problem that is caused by limited self-knowledge. Providing incentives for someone who is clueless simply ensures that he or she is both clueless and committed.

What, then, does mitigate CEO's problems with self-knowledge? One answer is, adding outside perspectives. It is not surprising that this works at the CEO level, given that many of the repairs that organizations develop for their employees involve forcing the employees to discuss decisions with other people who have an outside perspective (Heath et al., 1998). Hayward and Hambrick (1997) found that CEOs who scored high on their hubris measures were less likely to pay high acquisition premiums when there was an independent chairman of the board and when the board had a relatively high proportion of outsiders (i.e., directors who were not employees, suppliers, or buyers of the organization). Malmendier and Tate (2003) found that CEOs' tendency to acquire other firms is strongly mitigated by "effective corporate governance," defined as having a board of directors that is not too large (in the corporate governance literature, this is considered to be under 12 members, a number that social psychologists might still consider too high for effective group discussions).

Another way of mitigating the problems is by taking away CEOs' access to money to invest. Because CEOs think the market is wrong about their prospects, they typically regard outside financing as too expensive and prefer to finance their projects internally. If the internal cash is taken away, CEOs behave more responsibly because they have to convince outsiders to loan them money for new projects. Organizations have many routines and procedures, such as loan limits, that repair front-line loan officers' tendency to be overly optimistic; a similar repair might be effective for CEOs.

It is important, however, that outsiders be able to maintain their outside perspective. Even "outside" advisors and consultants may be seduced into focusing on the inside view. Moore and Kim (2003) showed that when outsiders had the opportunity to bet on the outcomes of a particular individual, they made errors almost as large as that individual did. This finding suggests that outsiders may behave like insiders when they face common incentives. Interestingly, in Hayward and Hambrick's (1997) study, higher percentages of outside directors helped reduce acquisition premiums, but not the degree to which they held the firm's stock. This suggests that it is the outside perspective and not necessarily the incentives of the outsiders that improved judgment. Indeed, when officers and directors of the company held more stock, companies tended to pay higher premiums to make acquisitions, suggesting that when boards are highly motivated to increase the value of the firm, they may actually make worse decisions.

CONCLUDING REMARKS

It was more than two centuries ago that Benjamin Franklin observed just how hard it was to know one's self. The data we have reviewed suggest that this difficulty remains in current times. The views people hold of themselves are often flawed. The correlation between those views and their objective behavior is often meager to modest, and people often claim to have valuable skills and desirable attributes to a degree that they do not.

However, we again hasten to add that gaining an accurate impression of the self is an intrinsically difficult task, one for which people often do not have crucial information. Thus, omniscience is a rare or impossible commodity, and one should not expect it of people, nor chide them for their failures to possess it. That said, people could take steps to reach more veridical conclusions about themselves. People should take into account more cautiously what they know and do not know about themselves and so adjust their self-views and predictions accordingly.

It is important to note that researchers have observed mistaken self-judgments not only in laboratory experiments, but also, more important, in real-world settings. Students have a difficult time understanding when they have grasped the true meaning of material they have just read. Surgical residents cannot predict how they will perform on a standardized exam on their skill, even though their peers and supervisors can. Patients act on models of illness that are erroneous and potentially harmful. College students, thinking themselves deviant in their ambivalence about alcohol, drink more than they would prefer to conform to a social norm that does not really exist. Planners of large civic projects envision completion schedules that bear little relation to reality. CEOs make acquisition decisions that the market punishes rather than rewards.

All told, this review suggests that there is striking continuity in the errors that people make when assessing themselves, whether in the laboratory or the real world. The review also suggests continuity in the psychological processes underlying these errors.

Recommendations

This review of the literature suggests many recommendations, some relevant to researchers and some to policymakers and practitioners.

For Researchers

Focusing on Consequences. In this review, we have concentrated on the types of errors people make in their self-judgments and the psychological processes that are responsible for those errors. Along the way, we have enumerated some of the consequences of these errors, but our discussion of this topic has not been as extensive or systematic. In large part, this is because behavioral scientists know a good deal about the types of errors people make

and the psychology that produces them, but much less about the sequelae of those errors.

To be sure, the literature is not bereft of discussion about consequences. Studies have provided a laundry list of circumstances in which self-judgment errors, particularly overconfidence, are either helpful or harmful. For example, undue optimism appears to be helpful when people encounter the most stressful of psychological circumstances. Women facing breast cancer display the most psychological adjustment when they exhibit unrealistic levels of optimism (Taylor, Lichtman, & Wood, 1984). Similarly, people recovering from civil war or the death of a loved one tend to do better when they overrate themselves (Bonnanno, Field, Kovacevic, & Kaltman, 2002). However, in less extreme circumstances, overconfidence appears not to be so beneficial. People who overrate themselves tend to be seen as arrogant, hostile, and maladjusted in the eyes of others (Colvin, Block, & Funder, 1995; Paulhus, 1998). Overconfident people persist in working on insolvable problems longer than people not so confident (Feather, 1968); overconfident business owners persist in outmoded strategies when economic conditions change (Audia, Locke, & Smith, 2000). Poor-performing business managers who overrate themselves are the ones most likely to be derailed in their careers (Shipper & Dillard, 2000).

Much future work could profitably focus more on the consequences of mistaken self-assessments, to provide a more comprehensive and systematic account of when mistaken judgment is likely to produce its greatest costs, as well as when it might provide valuable benefits. Such work would have to examine carefully the impact of different types of self-judgment errors, the magnitude of those errors, and the settings in which those errors take place (for a more extensive discussion, see Dunning, 2005).

Creating a Unified Research Focus. In addition, the continuity of the patterning of self-judgment errors in different domains suggests that the fallibility of self-evaluation is a ubiquitous issue that arises across disparate human endeavors. However, it appears that up to now, researchers have been largely content to study the accuracy of self-judgment within the confines of their own specific specialty, not realizing that other researchers in other specialties have found similar patterns of self-evaluation strengths and weaknesses. For example, some research on CEO hubris does not cite the extensive literature on the psychology of overconfidence. The psychology literature on whether overconfidence is merely a laboratory epiphenomenon could benefit from a consideration of the evidence that similar errors are encountered in deadly serious real-life circumstances in the medical field. The medical literature does not cite organizational research about when people are most likely to be overconfident (e.g., perhaps like CEOs, doctors are especially prone to misguided confidence when they confront the new).

This continuity suggests that researchers across a number of disciplines and subdisciplines need to make the accuracy of self-judgment an explicit, identified topic of empirical research and to

interact with one another, or at least track each other's work. Historically, this type of interdisciplinary cross talk has been somewhat unusual, but with the advent of electronic databases, the ability of medical researchers, for example, to reference the most recent findings in the organizational or educational literatures has been made much easier. By making such references, researchers in one subdiscipline may be able to find out what precious information other disciplines have uncovered about when people are likely to be right or wrong in their self-appraisals, what psychological mechanisms produce those patterns of accuracy or error, and what interventions might bring self-perception into a closer alignment with reality.

For Policymakers and Practitioners

For policymakers and real-world practitioners, the implication of this review is that the accuracy of self-evaluation should not be assumed. For example, in business settings, one should not assume that employees have achieved the level of expertise that they claim. Instead, one should take pains to provide independent tests of competence (such as the opinions of other people). We have enumerated a number of flaws typical of self-evaluation, and practitioners should be wary of how these types of flaws might be relevant to their own work.

For example, consider the area of medical education. Many medical schools emphasize to their students that they should independently develop the initiative, habits, and expertise necessary to educate themselves about the types of situations and challenges they will face in the classroom, the medical-school clinic, and, ultimately, in their own offices after beginning their practice (Brockett & Hiemstra, 1991; Candy, 1991). Obviously, such self-directed learning requires the ability to recognize the areas where further work is most needed—where one's shortcomings are the most severe and in need of remediation. This review, however, suggests that one cannot simply assume that individuals, left to their own devices, will be able to spot their own shortcomings. Therefore, it might be prudent to provide some sort of intervention (e.g., peer review, standard tests) that gives students—and practitioners after they have left formal training—feedback about their strengths and weaknesses.

Some common themes run through the literature on improving the accuracy of self-judgment. One theme that emerges from our review is that the road to self-accuracy may involve information from or about other people (see Dunning, 2005, for an extended discussion). For example, in educational settings, benchmarking has been shown to improve self-evaluation accuracy, as has peer assessment. In the business world, having an independent and active board of directors has been shown to prevent CEOs from making the kind of mistakes that grow out of hubris. Another theme, coming from the organizational literature, is that cognitive repairs can be applied to the kinds of self-judgments that are often made with error, thus sparing individuals and their organizations the costs associated with faulty self-assessment.

Unfinished Portraits

In any event, this review of the literature has prompted us to believe that for both researcher and practitioner, there is much exciting work to be done on the psychology of faulty self-assessment. We feel that the psychological literature has painted only a few brushstrokes toward a portrait of the person as self-evaluator—and there is much more painting to be done to complete that portrait. But, perhaps more important, there is also much work to be done about another portrait well worth painting. That second portrait is one that depicts what an individual looks like when he or she has achieved an accurate impression of his or her talents, capacities, and character. How one retouches the first portrait to create the second is an issue that requires much more theoretical and empirical work.

Acknowledgments—The writing of this article was supported financially by National Institute of Mental Health Grant RO1 56072 to David Dunning and by National Science Foundation Grant BCS-99-10592 to Jerry M. Suls.

REFERENCES

- Agostinelli, G., Brown, J.M., & Miller, W.R. (1995). Effects of normative feedback on consumption among heavy drinking college students. *Journal of Drug Education, 25*, 31–40.
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Alicke, M.D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49*, 1621–1630.
- Amler, R., Moriarty, D., & Hutchins, E. (Eds.). (1989). *Healthier people*. Decatur, GA: Carter Center of Emory University Health Risk Appraisal Program.
- Arnold, L., Willoughby, T.L., & Calkins, E.V. (1985). Self-evaluation in undergraduate medical education: A longitudinal perspective. *Journal of Medical Education, 60*, 21–28.
- Ashford, S.J. (1989). Self-assessments in organizations: A literature review and integrative model. In B.M. Staw & L.L. Cummings (Eds.), *Research in organizational behavior* (Vol. 11, pp. 133–174). Greenwich, CT: JAI Press.
- Audia, P.G., Locke, E.A., & Smith, K.G. (2000). The paradox of success: An archival and a laboratory study of strategic persistence following radical environmental change. *Academy of Management Journal, 43*, 837–853.
- Baddeley, A.D., & Longman, D.J.A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21*, 627–635.
- Barney, J.B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*, 99–120.
- Baron, J.N., & Kreps, D.M. (1999). *Strategic human resources*. New York: John Wiley & Sons.
- Barrows, H.S., & Tamblyn, R.M. (1980). *Problem-based learning: An approach to medical education*. New York: Springer.
- Bass, B.M., & Yammarino, F.J. (1991). Congruence of self and others' leadership ratings of Naval officers for understanding successful performance. *Applied Psychology, 40*, 437–454.
- Baumann, L., Cameron, L., Zimmerman, R., & Leventhal, H. (1989). Illness representations and matching labels with symptoms. *Health Psychology, 8*, 449–469.

- Baumann, L., & Leventhal, L. (1985). "I can tell when my blood pressure is up, can't I?" *Health Psychology*, 4, 203–218.
- Baumeister, R.F. (1990). The optimal margin of illusion. *Journal of Social and Clinical Psychology*, 8, 176–189.
- Baumeister, R.F., & Newman, L.S. (1994). Self-regulation of cognitive inference and decision processes. *Personality and Social Psychology Bulletin*, 20, 3–19.
- Becker, M.H. (1974). The health belief model and personal health behavior. *Health Education Monographs*, 2, 329–508.
- Belloc, N.B. (1973). Relationship of health practices and mortality. *Preventive Medicine*, 2, 67–81.
- Benjamin, A.S., Bjork, R.A., & Schwartz, B.L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, 127, 55–68.
- Bjork, R.A. (1988). Retrieval practice and the maintenance of knowledge. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical aspects of memory: Current research and issues: Vol. 1. Memory in everyday life* (pp. 396–401). New York: John Wiley & Sons.
- Bjork, R.A. (1994). Institutional impediments to effective training. In D. Druckman & R.A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 295–306). Washington, DC: National Academy Press.
- Bjork, R.A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriati (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 435–459). Cambridge, MA: MIT Press.
- Bol, L., & Hacker, D.J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *Journal of Experimental Education*, 69, 133–151.
- Bonanno, G.A., Field, N.P., Kovacevic, A., & Kaltman, S. (2002). Self-enhancement as a buffer against extreme adversity: Civil war in Bosnia and traumatic loss in the United States. *Personality and Social Psychology Bulletin*, 28, 184–196.
- Borkenau, P., & Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology*, 65, 546–553.
- Borsari, B., & Carey, K. (2001). Peer influences on college drinking: A review of the research. *Journal of Substance Abuse*, 13, 391–424.
- Boud, D. (1995). *Enhancing learning through self assessment*. Philadelphia: Kogan Page.
- Boud, D., & Feletti, G. (1991). *The challenge of problem-based learning*. Philadelphia: Kogan Page.
- Bourgeois, M., & Bowen, A. (2001). Self-organization of alcohol-related attitudes and beliefs in a campus housing complex: An initial investigation. *Health Psychology*, 20, 434–438.
- Braden, B.R., Cowan, C.A., Lazenby, H.C., Martin, A.B., McDonnell, P.A., Sensenig, A.L., Stiller, J.M., Whittle, L.S., Donham, C.S., Long, A.M., & Stewart, M.W. (1998). National health expenditures in 1997. *Health Care Financing Review*, 20, 83–126.
- Brockett, R.G., & Hiemstra, R. (1991). *Self-direction in adult learning: Perspectives on theory, research, and practice*. London: Routledge.
- Brody, J.E. (1999, April 27). Chronic heartburn, an ominous warning. *New York Times*, p. D6.
- Buehler, R., Griffin, D., & MacDonald, H. (1997). The role of motivated reasoning in optimistic time predictions. *Personality and Social Psychology Bulletin*, 23, 238–247.
- Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67, 366–381.
- Buehler, R., Griffin, D., & Ross, M. (2002). Inside the planning fallacy: The causes and consequences of optimistic time predictions. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 251–270). Cambridge, England: Cambridge University Press.
- Burger, J.M., & Burns, L. (1988). The illusion of unique invulnerability and the use of effective contraception. *Personality and Social Psychology Bulletin*, 14, 264–270.
- Burger, J.M., & Palmer, M. (1992). Changes in and generalization of unrealistic optimism following experiences from stressful events: Reactions to the 1989 California earthquake. *Personality and Social Psychology Bulletin*, 18, 39–43.
- Buunk, B., Gibbons, F., & Visser, A. (2002). The relevance of social comparison processes for prevention and health care. *Patient Education and Counseling*, 47, 1–3.
- Cambridge, R.M., & Shreckengost, R.C. (1980). *Are you sure? The subjective probability assessment test*. Unpublished manuscript, Office of Training, Central Intelligence Agency, Langley, VA.
- Camerer, C., & Lavallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89, 306–318.
- Cameron, L.E., Leventhal, E., & Leventhal, H. (1995). Seeking medical care in response to symptoms and stressors. *Psychosomatic Medicine*, 57, 37–47.
- Candy, P.C. (1991). *Self-direction for lifelong learning: A comprehensive guide to theory and practice*. San Francisco: Jossey-Bass.
- Caputo, D., & Dunning, D. (in press). What you do not know: The role played by errors of omission in imperfect self-assessment. *Journal of Experimental Social Psychology*.
- Chambers, J., Windschitl, P., & Suls, J. (2003). Egocentrism, event frequency and comparative optimism: When what happens frequently is "More likely to happen to me." *Personality and Social Psychology Bulletin*, 29, 1343–1356.
- Chassin, L., Presson, C., Sherman, J., Corty, E., & Olshavsky, R. (1984). Predicting the onset of cigarette smoking in adolescents: A longitudinal study. *Journal of Applied Social Psychology*, 14, 224–243.
- Chemers, M.M., Hu, L., & Garcia, B.F. (2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology*, 93, 55–64.
- Chi, M.T.H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 17–76). Hillsdale, NJ: Erlbaum.
- Christensen, P., & Glad, A. (1996). Mandatory course of driving on slippery roads does not reduce the accident rate. *Nordic Road and Transport Research*, 8, 22–23.
- Christiansen-Szalanski, J.J.J., & Bushyhead, J.B. (1981). Physicians' use of probabilistic information in a real clinical setting. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 928–935.
- Clark, M. (1959). *Health in the Mexican-American culture*. Berkeley: University of California Press.
- Cochrane, S.B., & Spears, M.C. (1980). Self-assessment and instructors' ratings: A comparison. *Journal of the American Dietetic Association*, 76, 253–255.
- College Board. (1976–1977). *Student descriptive questionnaire*. Princeton, NJ: Educational Testing Service.
- Colvin, C.R., & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown formulation. *Psychological Bulletin*, 116, 3–20.
- Colvin, C.R., Block, J., & Funder, D. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology*, 68, 1152–1162.

- Cooper, A.C., Woo, C.Y., & Dunkelberg, W.C. (1988). Entrepreneurs' perceived chances for success. *Journal of Business Venturing*, 3, 97–108.
- Crosby, R.A., & Yarber, W.L. (2001). Perceived versus actual knowledge about correct condom use among U.S. adolescents. *Journal of Adolescent Health*, 28, 415–420.
- Cross, P. (1977). Not can but *will* college teaching be improved? *New Directions for Higher Education*, 17, 1–15.
- Cusumano, M.A., & Selby, R.W. (1995). *Microsoft secrets*. New York: Free Press.
- Dancer, W.T., & Dancer, J. (1992). Peer rating in higher education. *Journal of Education for Business*, 67, 306–309.
- Davis, D. (1985, March–April). New projects: Beware of false economies. *Harvard Business Review*, 63, 95–101.
- Dawes, R.M. (1988). *Rational choice in an uncertain world*. San Diego, CA: Harcourt Brace Jovanovich.
- DeMarco, T. (1982). *Controlling software projects: Management, measurement, and estimation*. New York: Yourdon.
- Dempster, F.N. (1990). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist*, 43, 627–634.
- DeNisi, A.S., & Kluger, A.N. (2000). Feedback effectiveness: Can 360-degree appraisals be improved? *Academy of Management Executive*, 14, 129–139.
- DePaulo, B.M., Charlton, K., Cooper, H., Lindsay, J.J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review*, 1, 346–357.
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24, 331–350.
- Dracup, K., Alonzo, A.A., Atkins, J.M., Bennett, N.M., Braslow, A., & Clark, L.T., et al. (1997). The physician's role in minimizing pre-hospital delay in patients at high risk of acute myocardial infarction: Recommendations from the National Heart Attack Alert Program. *Annals of Internal Medicine*, 126, 645–651.
- Duch, B., Gron, S., & Allen, D. (2001). *The power of problem-based learning*. Sterling, VA: Stylus.
- Dufresne, A., & Kobasigawa, A. (1988). Developmental differences in children's spontaneous allocation of study time. *Journal of Genetic Psychology*, 149, 87–92.
- Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology*, 47, 274–296.
- Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J. Sternberg & J. Davidson (Eds.), *The nature of insight* (pp. 365–395). Cambridge, MA: MIT Press.
- Dunlosky, J., & Nelson, T.O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, 20, 374–380.
- Dunning, D. (2001). On the motives underlying social cognition. In N. Schwarz & A. Tesser (Eds.), *Blackwell handbook of social psychology: Vol. 1. Intraindividual processes* (pp. 348–374). New York: Blackwell.
- Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychology Press.
- Dunning, D. (in press). Prediction: The inside view. In A. Kruglanski & E. Higgins (Eds.), *Basic principles in social psychology*. New York: Guilford.
- Dunning, D., Griffin, D.W., Milojkovic, J.H., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, 58, 568–592.
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12, 83–86.
- Dunning, D., Meyerowitz, J.A., & Holzberg, A.D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, 57, 1082–1090.
- Dunning, D., Perie, M., & Story, A.L. (1991). Self-serving prototypes of social categories. *Journal of Personality and Social Psychology*, 61, 957–968.
- Dunning, D., & Story, A.L. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? *Journal of Personality and Social Psychology*, 61, 521–532.
- Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H.A. Ruger & C.E. Bussenius, Trans.). New York: Dover. (Original work published 1885)
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2004). *Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent*. Unpublished manuscript, Cornell University, Ithaca, NY.
- Epley, N., & Dunning, D. (2000). Feeling “holier than thou”: Are self-serving assessments produced by errors in self or social prediction? *Journal of Personality and Social Psychology*, 79, 861–875.
- Epley, N., & Dunning, D. (2004). *The mixed blessings of self-knowledge in behavioral prediction*. Unpublished manuscript, Harvard University, Cambridge, MA.
- Epstein, W., Glenberg, A.M., & Bradley, M. (1984). Coactivation and comprehension: Contribution of text variables to the illusion of knowing. *Memory & Cognition*, 12, 355–360.
- Fagerlin, A., Ditto, P., Danks, J., Houts, R., & Smucker, W. (2001). Projection in surrogate decisions about life-sustaining medical treatments. *Health Psychology*, 20, 166–175.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59, 395–430.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70, 287–322.
- Farh, J.L., & Dobbins, G.H. (1989). Effects of comparative performance information on the accuracy of self-ratings and agreement between self- and supervisor ratings. *Journal of Applied Psychology*, 74, 606–610.
- Feather, N.T. (1968). Change in confidence following success or failure as a predictor of subsequent performance. *Journal of Personality and Social Psychology*, 9, 38–46.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 552–564.
- Fouad, N.A., Smith, P.L., & Zao, K.E. (2002). Across academic domains: Extensions of the social-cognitive career model. *Journal of Counseling Psychology*, 49, 164–171.
- Friedman, M., & Rosenman, R. (1959). *Type A and your heart*. New York: Knopf.
- Friedrich, J. (1996). On seeing oneself as less self-serving than others: The ultimate self-serving bias? *Teaching of Psychology*, 23, 107–109.

- Friend, K., & Koushki, P. (1984). Student substance abuse: Stability and change across college years. *Public Opinion Quarterly*, 40, 427–448.
- Funder, D.C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin*, 101, 75–90.
- Gibbons, F., & Gerrard, M. (1995). Predicting young adults' health risk behavior. *Journal of Personality and Social Psychology*, 69, 505–517.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review*, 103, 592–596.
- Gigerenzer, G., Hoffrage, U., & Kleinboelting, H. (1991). Probabilistic mental models: A Brunswickian theory of confidence. *Psychological Review*, 98, 506–528.
- Gilbert, D.T., Pinel, E.C., Wilson, T.D., Blumberg, S.J., & Wheatley, T.P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 75, 617–638.
- Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.
- Glenberg, A.M. (1979). Component-level theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, 7, 95–112.
- Glenberg, A.M. (1992). Distributed practice effects. In L.R. Squire (Ed.), *Encyclopedia of learning and memory* (pp. 138–142). New York: Macmillan.
- Glenberg, A.M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition*, 15, 84–93.
- Glenberg, A.M., Wilkinson, A., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition*, 10, 597–602.
- Gold, R., & Aucote, H. (2003). "I'm less at risk than most guys": Gay men's unrealistic optimism about becoming infected with HIV. *International Journal of STD and AIDS*, 14, 18–23.
- Griffin, D.W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predictions about the self and others. *Journal of Personality and Social Psychology*, 59, 1128–1139.
- Griffin, D.W., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 319–359). San Diego, CA: Academic Press.
- Guest, C.B., Regehr, G., & Tiberius, R.G. (2001). The life long challenge of expertise. *Medical Education*, 35, 78–81.
- Hacker, D.J., Bol, L., Horgan, D., & Rakow, E. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, 92, 160–170.
- Hall, K., Domingues, D., & Cavazos, R. (1994). The effects of contextual interference on college level baseball players. *Perceptual and Motor Skills*, 78, 835–841.
- Hansen, W., & Graham, J. (1991). Preventing alcohol, marijuana and cigarette use among adolescents: Peer resistance training versus establishing conservative norms. *Preventive Medicine*, 20, 414–430.
- Hansford, B.C., & Hattie, J.A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research*, 52, 123–142.
- Harris, M.M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43–62.
- Harrison, J.A., Mullen, P.D., & Green, L.W. (1992). A meta-analysis of studies of the health belief model with adults. *Health Education Review*, 7, 107–116.
- Hastie, R., & Dawes, R.M. (2001). *Rational choice in an uncertain world*. Thousand Oaks, CA: Sage.
- Haun, D.E., Zeringue, A., Leach, A., & Foley, A. (2000). Assessing the competence of specimen-processing personnel. *Laboratory Medicine*, 31, 633–637.
- Hayes, A.F., & Dunning, D. (1997). Construal processes and trait ambiguity: Implications for self-peer agreement in personality judgment. *Journal of Personality and Social Psychology*, 72, 664–677.
- Hayward, M.L.A., & Hambrick, D.C. (1997). Explaining the premiums paid for large acquisitions: Evidence of CEO hubris. *Administrative Science Quarterly*, 42, 103–127.
- Heath, C., Larrick, R.P., & Klayman, J. (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. In B.M. Staw & L.L. Cummings (Eds.), *Research in organizational behavior*, Vol. 20 (pp. 1–37). Greenwood, NJ: JAI Press.
- Heaton, J.B. (2002). Managerial optimism and corporate finance. *Financial Management*, 31, 33–45.
- Heine, S.J., & Lehman, D.R. (1995). Cultural variation in unrealistic optimism: Does the West feel more vulnerable than the East? *Journal of Personality and Social Psychology*, 68, 595–607.
- Heine, S.J., Lehman, D.R., Markus, H.R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, 106, 766–794.
- Helweg-Larsen, M., & Shepperd, J. (2001). Do moderators of the optimistic bias affect personal or target risk estimates? A review of the literature. *Personality and Social Psychology Review*, 5, 74–95.
- Hines, D., Saris, R.N., & Throckmorton-Belzer, L. (2002). Pluralistic ignorance and health risk behaviors: Do college students misperceive social approval for risky behaviors on campus and in media? *Journal of Applied Social Psychology*, 32, 2621–2640.
- Hodges, B., Regehr, G., & Martin, D. (2001). Difficulties in recognizing one's own incompetence: Novice physicians who are unskilled and unaware of it. *Academic Medicine*, 76, S87–S89.
- Jacoby, L.L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, 17, 649–667.
- Jemmott, J., Croyle, R., & Ditto, P. (1986). Judging health status: Effects of perceived prevalence and personal relevance. *Journal of Personality and Social Psychology*, 50, 899–905.
- Ji, L., Zhang, Z., Osborne, E., & Guan, Y. (2004). Optimism across cultures: In response to the severe acute respiratory syndrome outbreak. *Asian Journal of Social Psychology*, 7, 25–34.
- Johnson, J., & King, K. (1995). Influence of expectations about symptoms on delay in seeking treatment during a myocardial infarction. *American Journal of Critical Care*, 4, 29–35.
- Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39, 17–31.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. *TIMS Studies in Management Science*, 12, 313–327.
- Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 509–520). Cambridge, England: Cambridge University Press.

- Kelemen, W.L. (2000). Metamemory cues and monitoring accuracy: Judging what you know and what you will know. *Journal of Educational Psychology*, 92, 800–810.
- Kelemen, W.L., & Weaver, C.A. III. (1997). Enhanced metamemory at delays: Why do judgments of learning improve over time? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1394–1409.
- Kelley, H. (1967). Attribution theory in social psychology. In D.L. Vine (Ed.), *Nebraska symposium on motivation* (Vol. 15, pp. 192–241). Lincoln: University of Nebraska Press.
- Keren, G.B. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes*, 139, 98–114.
- Kim, Y.M., Putjuk, F., Kols, A., & Basuki, E. (2000). *Improving provider-client communication: Reinforcing IPC/C training in Indonesia with self-assessment and peer review*. Bethesda, MD: Quality Assurance Project.
- Klein, C., & Helweg-Larsen, M. (2002). Perceived control and the optimistic bias: A meta-analytic review. *Psychology and Health*, 17, 437–446.
- Klein, W.M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology*, 72, 763–774.
- Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284.
- Koch, A. (2001). Training in metacognition and comprehension of physics texts. *Science Education*, 75, 858–868.
- Kolar, D.W., Funder, D.C., & Colvin, C.R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality*, 64, 331–337.
- Kreuter, M.W., & Strecher, V.J. (1995). Changing inaccurate perceptions of health risk: Results from a randomized trial. *Health Psychology*, 14, 56–63.
- Krueger, J., & Mueller, R.A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology*, 82, 180–188.
- Kruger, J. (1999). Lake Wobegon be gone! The “below-average effect” and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology*, 77, 221–232.
- Kruger, J., & Burrus, J. (2004). Egocentrism and focalism in unrealistic optimism (and pessimism). *Journal of Experimental Social Psychology*, 40, 332–340.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77, 1121–1134.
- Kruger, J., & Dunning, D. (2002). Unskilled and unaware—But why? A reply to Krueger and Mueller. *Journal of Personality and Social Psychology*, 82, 189–192.
- Kulik, J., & Mahler, H. (1987). Health status, perceptions of risk and prevention interest for health and nonhealth problems. *Health Psychology*, 6, 15–27.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498.
- Lagergren, J., Bergstrom, R., Lindgren, A., & Nyren, O. (1999). Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. *New England Journal of Medicine*, 340, 825–831.
- Lambert, T.A., Kahn, A.S., & Apple, K.J. (2003). Pluralistic ignorance and hooking up. *Journal of Sex Research*, 40, 129–133.
- Landier, A., & Thesmar, D. (2003). *Financial contracting with optimistic entrepreneurs: Theory and evidence*. Unpublished manuscript, University of Chicago, Chicago.
- Lane, J.L., & Gottlieb, R.P. (2004). Improving the interviewing and self-assessment skills of medical students: Is it time to readopt videotaping as an educational tool? *Ambulatory Pediatrics*, 4, 244–248.
- Larson, J.R. Jr. (1989). The dynamic interplay between employees’ feedback-seeking strategies and supervisors’ delivery of performance feedback. *Academy of Management Review*, 14, 408–422.
- Larwood, L. (1978). Swine flu: A field study of self-serving biases. *Journal of Applied Social Psychology*, 8, 283–289.
- Larwood, L., & Whittaker, W. (1977). Managerial myopia: Self-serving biases in organizational planning. *Journal of Applied Psychology*, 62, 194–198.
- Lennon, S. (1995). Correlations between tutor, peer and self assessments of second year physiotherapy students in movement studies. In S. Griffiths, K. Houston, & A. Lazenblatt (Eds.), *Enhancing student learning through peer tutoring in higher education: Section 3—Implementing* (Vol. 1, pp. 66–71). Coleraine, Northern Ireland: University of Ulster.
- Lent, R.W., Brown, S.D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45, 79–122.
- Leventhal, H., Meyer, D., & Nerenz, D. (1980). The common sense representation of illness danger. In S. Rachman (Ed.), *Contributions to medical psychology* (Vol. 2, pp. 7–30). New York: Pergamon.
- Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, 75, 5–18.
- Loftus, E.F., & Wagenaar, W.A. (1988). Lawyers’ predictions of success. *Jurimetrics Journal*, 29, 437–453.
- Lovallo, D., & Kahneman, D. (2003, July). Delusions of success: How optimism undermines executives’ decision. *Harvard Business Review*, 81, 56–63.
- Mabe, P.A. III, & West, S.G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67, 280–286.
- MacDonald, T.K., & Ross, M. (1999). Assessing the accuracy of predictions about dating relationships: How and why do lovers’ predictions differ from those made by observers? *Personality and Social Psychology Bulletin*, 25, 1417–1429.
- Madey, S., & Gomez, R. (2003). Reduced optimism for perceived age-related medical conditions. *Basic and Applied Social Psychology*, 25, 213–219.
- Maki, R.H. (1998). Test predictions over text material. In D.J. Hacker, J. Dunlosky, & A.C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117–144). Mahwah, NJ: Erlbaum.
- Maki, R.H., & Berry, S.L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 663–679.
- Maki, R.H., Jonas, D., & Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin & Review*, 1, 126–129.
- Malmendier, U., & Tate, G. (2003). *Who makes acquisitions? A test of the overconfidence hypothesis* (Stanford Research Paper 1798). Stanford, CA: Stanford University.

- Malmendier, U., & Tate, G. (in press). CEO overconfidence and corporate investment. *Journal of Finance*.
- Marcoulides, G., & Simkin, M.G. (1991). Evaluating student papers: The case for peer review. *Journal of Education for Business*, 67, 80–83.
- Marteau, T.M., Johnston, M., Wynne, G., & Evans, T.R. (1989). Cognitive factors in the explanation of the mismatch between confidence and competence in performing basic life support. *Psychology and Health*, 3, 173–182.
- Martin, D., Regehr, G., Hodges, B., & McNaughton, N. (1998). Using videotaped benchmarks to improve the self-assessment ability of family practice residents. *Academic Medicine*, 73, 1201–1206.
- Martin, R., Gordon, E., & Lounsbury, P. (1998). Gender disparities in the interpretation of cardiac-related symptoms: The contribution of common-sense models of illness. *Health Psychology*, 17, 346–357.
- Martin, R., & Lemos, K. (2002). From heart attacks to melanoma: Do common sense models of somatization influence symptom interpretation for female victims? *Health Psychology*, 21, 25–32.
- Martin, R., Lemos, K., Rothrock, N., Bellman, S.B., Russell, D., Tripp-Reimer, T., Lounsbury, P., & Gordon, E. (2004). Gender disparities in common sense models of illness among myocardial infarction victims. *Health Psychology*, 23, 345–353.
- Martin, R., & Leventhal, H. (2004). Symptom perception and health care-seeking behavior. In T. Boll, J. Raczynski, & L. Leviton (Eds.), *Handbook of clinical health psychology* (Vol. 2, pp. 299–328). Washington, DC: American Psychological Association.
- Mayhew, D.R., & Simpson, H.M. (1996). *Effectiveness and role of driver education and training in a graduated licensing system*. Ottawa, Ontario, Canada: Traffic Injury Research Foundation.
- McKenna, F., & Albery, I. (2001). Does unrealistic optimism change following a negative experience? *Journal of Applied Social Psychology*, 31, 1146–1157.
- McKenna, F.P. (1993). It won't happen to me: Unrealistic optimism or illusion of control? *British Journal of Psychology*, 84, 39–50.
- Mechanic, D. (1972). Social psychological factors affecting the presentation of bodily complaints. *New England Journal of Medicine*, 286, 1132–1139.
- Meindl, J.R., Ehrlich, S.R., & Dukerich, J.M. (1985). The romance of leadership. *Administrative Science Quarterly*, 30, 78–102.
- Meyer, D., Leventhal, H., & Gutmann, M. (1985). Common-sense models of illness: The example of hypertension. *Health Psychology*, 4, 115–135.
- Middleton, W., Harris, P., & Surman, M. (1996). Give 'em enough rope: Perception of health and safety risks in bungee jumpers. *Journal of Social and Clinical Psychology*, 15, 68–79.
- Miller, D.T., & McFarland, C. (1987). Pluralistic ignorance: When similarity is interpreted as dissimilarity. *Journal of Personality and Social Psychology*, 53, 298–305.
- Miller, D.T., & McFarland, C. (1991). When social comparison goes awry: The case of pluralistic ignorance. In J. Suls & T.A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 287–313). Hillsdale, NJ: Erlbaum.
- Miller, D.T., & Nelson, L.D. (2002). Seeing approach motivation in the avoidance behavior of others: Implications for an understanding of pluralistic ignorance. *Journal of Personality and Social Psychology*, 83, 1066–1075.
- Mintzberg, H. (1975, July–August). The manager's job: Folklore and fact. *Harvard Business Review*, 53, 49–61.
- Moore, D.A., & Cain, D.M. (2004). *Strategic myopia and entrepreneurial entry: An experimental approach*. Unpublished manuscript, Carnegie Mellon University, Pittsburgh, PA.
- Moore, D.A., & Kim, T.G. (2003). Myopic social prediction and the solo comparison effect. *Journal of Personality and Social Psychology*, 85, 1121–1135.
- Moreland, R.L., Miller, J., & Laucka, F. (1981). Academic achievement and self-evaluations of academic performance. *Journal of Educational Psychology*, 73, 335–344.
- Morrison, V., Ager, A., & Willock, J. (1999). Perceived risk of tropical diseases in Malawi: Evidence of unrealistic pessimism and the irrelevance of beliefs of personal control? *Psychology, Health and Medicine*, 4, 361–368.
- Mullen, B., Atkins, J., Champion, D., Edwards, C., Hardy, D., Story, J., & Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology*, 21, 262–283.
- Mullen, B., & Riordan, C.A. (1988). Self-serving attributions for performance in naturalistic settings: A meta-analytic review. *Journal of Applied Social Psychology*, 18, 3–22.
- Nelson, T.O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The delayed-JOL effect. *Psychological Science*, 2, 267–270.
- Newby-Clark, I.R., Ross, M., Buehler, R., Koehler, D.J., & Griffin, D. (2000). People focus on optimistic and disregard pessimistic scenarios while predicting their task completion times. *Journal of Experimental Psychology: Applied*, 6, 171–182.
- Newell, A. (1969). Heuristic programming: Ill structured problems. In J. Aronofsky (Ed.), *Progress in operations research* (Vol. 3, pp. 360–414). New York: Wiley.
- Nisbett, R.E., & Ross, L.D. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Norcini, J. (2003). Peer assessment of competence. *Medical Education*, 37, 539–543.
- Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *Journal of Finance*, 8, 1887–1934.
- Oksam, J., Kingma, J., & Klasen, H.J. (2000). Clinicians' recognition of 10 different types of distal radial fractures. *Perceptual and Motor Skills*, 91, 917–924.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Clinical Psychology*, 29, 261–265.
- Paulhus, D.L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing. *Journal of Personality and Social Psychology*, 74, 1197–1208.
- Pennebaker, J.W. (1982). *The psychology of physical symptoms*. New York: Springer-Verlag.
- Perkins, H.W. (2002). Social norms and the prevention of alcohol abuse in collegiate contexts. *Journal of Studies in Alcohol*, 14, 164–172.
- Perkins, H.W., & Berkowitz, A. (1986). Perceiving the community norms of alcohol use among students: Some research implications for campus alcohol education programming. *International Journal of the Addictions*, 21, 961–976.
- Petrie, K.J., Cameron, L.D., Ellis, C.J., Buick, D., & Weinman, J. (2002). Changing illness perceptions following myocardial infarction: An early intervention randomized controlled trial. *Psychosomatic Medicine*, 64, 580–586.
- Porter, M. (1980). *Competitive strategy*. New York: Free Press.
- Prahalad, C.K., & Hamel, G. (1990, May–June). The core competence of the corporation. *Harvard Business Review*, 68, 79–91.
- Prentice, D., & Miller, D.T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, 64, 243–256.

- Pronin, E., Lin, D.Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28, 369–381.
- Quadrel, M.J., Fischhoff, B., & Davis, W. (1993). Adolescent (in)vulnerability. *American Psychologist*, 48, 102–116.
- Read, D., & van Leeuwen, B. (1998). Time and desire: The effects of anticipated and experienced hunger and delay to consumption on the choice between healthy and unhealthy snack food. *Organizational Behavior and Human Decision Processes*, 76, 189–205.
- Reder, L.M., Charney, D.H., & Morgan, K.I. (1986). The role of elaborations in learning a skill from an instructional text. *Memory & Cognition*, 14, 64–78.
- Reed, M.B., & Aspinwall, L.G. (1998). Self-affirmation reduces biased processing of health risk information. *Motivation & Emotion*, 22, 99–132.
- Risucci, D.A., Torolani, A.J., & Ward, R.J. (1989). Ratings of surgical residents by self, supervisors and peers. *Surgical Gynecology and Obstetrics*, 169, 519–526.
- Rogers, R.W. (1975). A protection motivation theory of fear appraisals and attitude change. *Journal of Psychology*, 91, 93–114.
- Roll, R. (1986). The hubris hypothesis of corporate takeovers. *Journal of Business*, 59, 197–216.
- Rosenstock, I.M. (1966). Why people use health services. *Milbank Memorial Fund Quarterly*, 44, 94–124.
- Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13, 279–301.
- Rutter, D.R., Quine, L., & Albery, I.P. (1998). Perceptions of risk in motorcyclists: Unrealistic optimism, relative realism and predictions of behaviour. *British Journal of Psychology*, 89, 681–696.
- Saavedra, R., & Kwun, S.K. (1993). Peer evaluation in self-managing work groups. *Journal of Applied Psychology*, 78, 450–462.
- Sabini, J., Cosmas, K., Siepmann, M., & Stein, J. (1999). Underestimates and truly false consensus effects in estimates of embarrassment and other emotions. *Basic and Applied Social Psychology*, 21, 233–241.
- Sambell, K., & McDowell, L. (1997). The value of self- and peer assessment to the developing lifelong learner. In C. Rust (Ed.), *Improving student learning—improving students as learners* (pp. 56–66). Oxford, England: Oxford Centre for Staff and Learning Development.
- Schachter, S. (1951). Deviation, rejection and communication. *Journal of Abnormal and Social Psychology*, 46, 190–207.
- Schanck, R. (1932). A study of community and its group institutions conceived of as behavior of individuals. *Psychological Monographs*, 43(2), 1–133.
- Scherer, L.A., Chang, M.C., Meredith, J.W., & Battistella, F.D. (2003). Videotape review leads to rapid and sustained learning. *American Journal of Surgery*, 185, 516–520.
- Schkade, D.A., & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, 9, 340–346.
- Schneider, W., Vise, M., Lockl, K., & Nelson, T.O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning (JOL) task. *Cognitive Development*, 15, 115–134.
- Schroeder, C., & Prentice, D. (1998). Exposing pluralistic ignorance to reduce alcohol use among college students. *Journal of Applied Social Psychology*, 28, 2150–2180.
- Setterlund, M.B., & Niedenthal, P.M. (1993). "Who am I? Why am I here?": Self-esteem, self-clarity, and prototype matching. *Journal of Personality and Social Psychology*, 65, 769–780.
- Shaughnessy, J.J. (1979). Confidence judgment accuracy as a predictor of test performance. *Journal of Research in Personality*, 13, 505–514.
- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48, 1117–1131.
- Sheer, V.C., & Cline, R.J. (1994). The development and validation of a model explaining sexual behavior among college students: Implications for AIDS communication campaign. *Human Communication Research*, 21, 280–304.
- Sherman, D.A.K., Nelson, L.D., & Steele, C.M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin*, 26, 1046–1058.
- Sherman, J., Skov, R., Hervitz, E., & Stock, C. (1981). The effects of explaining hypothetical events: From possibility to probability to actuality and beyond. *Journal of Experimental Social Psychology*, 17, 142–158.
- Shipper, F., & Dillard, J.E. Jr. (2000). A study of impending derailment and recovery of middle managers across career stages. *Human Resource Management*, 39, 331–345.
- Shrauger, J.S., Ram, D., Greninger, S.A., & Mariano, E. (1996). Accuracy of self-predictions versus judgments by knowledgeable others. *Personality and Social Psychology Bulletin*, 22, 1229–1243.
- Sieff, E.M., Dawes, R.M., & Loewenstein, G. (1999). Anticipated versus actual reaction to HIV test results. *American Journal of Psychology*, 112, 297–311.
- Simon, D.A., & Bjork, R.A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 907–912.
- Simon, H.A. (1973). The structure of ill structured problems. *Artificial Intelligence*, 4, 181–201.
- Sinkavich, F.J. (1995). Performance and metamemory: Do students know what they don't know? *Instructional Psychology*, 22, 77–87.
- Skelly, G.B. (1968). *Aspects of driving experience in the first year as a qualified driver*. Crowthorne, England: Road Research Laboratory.
- Son, L.K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 204–221.
- Spain, J.S., Eaton, L.G., & Funder, D.C. (2000). Perspectives on personality: The relative accuracy of self versus others for the prediction of emotion and behavior. *Journal of Personality*, 68, 837–867.
- Sparks, P., & Shepherd, R. (1994). Public perceptions of the potential hazards associated with food production and food consumption: An empirical study. *Risk Analysis*, 14, 799–806.
- Stajkovic, A.D., & Luchins, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, 124, 240–261.
- Stock, M.F., Weaver, J.K., Ray, H.W., Brink, J.R., & Sadoff, M.G. (1983). *Evaluation of safe performance secondary school driver education curriculum demonstration project: Final report*. Washington, DC: U.S. Department of Transportation.
- Story, A.L. (2003). Similarity of trait and construal and consensus in interpersonal perception. *Journal of Experimental Social Psychology*, 39, 364–370.
- Stuart, M.R., Goldstein, H.S., & Snopce, F.C. (1980). Self-evaluation by residents in family medicine. *Journal of Family Practice*, 10, 639–642.
- Sullivan, M.E., Hitchcock, M.A., & Dunnington, G.L. (1999). Peer and self assessments during problem-based tutorials. *American Journal of Surgery*, 177, 266–269.

- Suls, J., & Green, P. (2003). Pluralistic ignorance and college student perceptions of gender-specific alcohol norms. *Health Psychology, 22*, 479–486.
- Suls, J., Lemos, K., & Stewart, H. (2003). Self-esteem, construal and comparisons with self, friends and peers. *Journal of Personality and Social Psychology, 82*, 252–261.
- Suls, J., & Wan, C.K. (1987). In search of the false uniqueness phenomenon: Fear and estimates of social consensus. *Journal of Personality and Social Psychology, 52*, 211–217.
- Suls, J., Wan, C.K., & Sanders, G. (1988). False consensus and false uniqueness in estimating the prevalence of health-protective behaviors. *Journal of Applied Social Psychology, 18*, 66–79.
- Suls, J., & Wheeler, L. (Eds.). (2000). *Handbook of social comparison*. New York: Kluwer.
- Sutton, S., & Bolling, K.D. (2003). Adolescents' perceptions of the risks and prevalence of smoking. *Zeitschrift für Gesundheitspsychologie, 11*, 107–111.
- Swim, J.K., & Hyers, L.L. (1999). Excuse me—what did you just say?!: Women's public and private responses to sexist remarks. *Journal of Experimental Social Psychology, 35*, 68–88.
- Taylor, S.E., & Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193–210.
- Taylor, S.E., & Brown, J.D. (1994). Positive illusions and well-being revisited: Separating fact from fiction. *Psychological Bulletin, 116*, 21–27.
- Taylor, S.E., Lichtman, R.R., & Wood, J.V. (1984). Attributions, beliefs about control, and adjustment to breast cancer. *Journal of Personality and Social Psychology, 46*, 489–502.
- Thiede, K.W. (1999). The importance of accurate monitoring and effective self-regulation during multitrial learning. *Psychonomic Bulletin & Review, 30*, 125–128.
- Thiede, K.W., & Anderson, M.C.M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology, 28*, 129–160.
- Thiede, K.W., Anderson, M.C.M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*, 66–73.
- Thiede, K.W., & Dunlosky, J. (1994). Delaying students' metacognitive monitoring improves their accuracy at predicting their recognition performance. *Journal of Educational Psychology, 86*, 290–302.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*, 249–276.
- Tracey, J.M., Arroll, B., Richmond, D.E., & Barham, P.M. (1997). The validity of general practitioners' self assessment of knowledge: Cross sectional study. *British Journal of Medicine, 315*, 1426–1428.
- Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110*, 403–422.
- Vallone, R.P., Griffin, D.W., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology, 58*, 582–592.
- Van Boven, L., Loewenstein, G., & Dunning, D. (in press). The illusion of courage in social predictions: Underestimating the impact of fear of embarrassment on other people. *Organizational Behavior and Human Decision Processes*.
- Van Boven, L., Loewenstein, G., Welch, E., & Dunning, D. (2004). *The illusion of courage: Underestimating fear of embarrassment in public performance decisions*. Unpublished manuscript, University of Colorado, Boulder.
- Vorauer, J.D., & Ratner, R.K. (1996). Who's going to make the first move? Pluralistic ignorance as an impediment to relationship formation. *Journal of Social and Personal Relationships, 13*, 483–506.
- Walczyk, J.J., & Hall, V.C. (1989). Effects of examples and embedded questions on the accuracy of comprehension self-assessments. *Journal of Educational Psychology, 81*, 435–437.
- Ward, M., MacRae, H., Schlachta, C., Mamazz, J., Poulin, E., Reznick, R., & Regehr, G. (2003). Resident self-assessment of operative performance. *American Journal of Surgery, 185*, 521–524.
- Wechsler, H., Davenport, A., Dowdall, G., Moeykens, B., & Castillo, S. (1994). Health and behavioral consequences of binge drinking in college: A national survey of students at 140 campuses. *Journal of the American Medical Association, 272*, 1672–1677.
- Weinstein, N.D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 38*, 806–820.
- Weinstein, N.D. (1983). Reducing unrealistic optimism about illness susceptibility. *Health Psychology, 2*, 11–20.
- Weinstein, N.D. (1987). Unrealistic optimism about susceptibility to health problems: Conclusions from a community-wide sample. *Journal of Behavioral Medicine, 10*, 481–500.
- Weinstein, N.D. (1988). The precaution adoption process. *Health Psychology, 7*, 355–386.
- Weinstein, N.D. (2003). Exploring the links between risk perceptions and preventive health behavior. In J. Suls & K. Wallston (Eds.), *Social psychological foundations of health and illness* (pp. 22–53). Malden, MA: Blackwell.
- Weinstein, N.D., & Klein, W. (1995). Resistance of personal risk perceptions to debiasing interventions. *Health Psychology, 14*, 132–140.
- Weinstein, N.D., & Lachendro, E. (1982). Egocentrism as a source of unrealistic optimism. *Personality and Social Psychology Bulletin, 8*, 195–200.
- Wilde, G.J.S. (1998). Targeting risk. *Recovery, 9*, 18–19.
- Wilson, T.D., Wheatley, T., Meyers, J.M., Gilbert, D.T., & Axson, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 78*, 821–836.
- Windschitl, P.D., Kruger, J., & Simms, E. (2003). The influence of egocentrism and focalism on people's optimism in competitions. *Journal of Personality and Social Psychology, 85*, 389–408.
- Woodzicka, J.A., & LaFrance, M. (2001). Real versus imagined gender harassment. *Journal of Social Issues, 57*, 15–30.
- Yates, J.F., Lee, J.W., & Bush, J.G. (1997). General knowledge overconfidence: Cross-national variations, response style, and "reality." *Organizational Behavior and Human Decision Processes, 70*, 87–94.
- Yates, J.F., Lee, J.W., Shinotsuka, H., Patalano, A.L., & Sieck, W.R. (1998). Cross-cultural variations in probability judgment accuracy: Beyond general knowledge overconfidence. *Organizational Behavior and Human Decision Processes, 74*, 89–117.
- Zenger, T.R. (1992). Why do employers only reward extreme performance? Examining the relationships among performance, pay, and turnover. *Administrative Science Quarterly, 37*, 198–219.