T. B. Gutkin & C. R. Reynolds (Eds.). (2009). Handbook of School Psychology (4th ed.) (pp. 210-229). Hoboken, NJ: Wiley.

CHAPTER 12

ERRORS IN DIAGNOSTIC DECISION MAKING AND CLINICAL JUDGMENT

MARLEY W. WATKINS
Arizona State University

Paradoxically, humans simultaneously attain extraordinary achievements and commit remarkable errors. Humans have walked on the moon, and died in the fiery ruins of space shuttles. Humans have extracted energy from atoms, and created a radioactive wasteland surrounding Chernobyl. Olympic athletes perform feats of strength and balance, but tourists stumble over guard rails and plummet into the Grand Canyon. Cognitive psychologists have speculated that this coincident capacity for attainment and error are "two sides of the same cognitive 'balance sheet' [where] each entry on the asset side carries a corresponding debit" (Reason, 1990, p. 2). For example, the lack of higher-level cognitive control during automatic performances allows smooth, highly integrated behavior but is vulnerable to distraction or preoccupation.

Given that errors are inevitable, it is crucial to identify when and how they might occur so that palliative action can be taken. Although there is no generally accepted taxonomy of error, Reason (1990) has articulated a tripartite generic error-modeling system that may be applied to school psychology.

GENERIC ERROR-MODELING SYSTEM

Skill-Based Errors

Behavior at the skill level is "primarily a way of dealing with routine and nonproblematic activities in familiar situations" (Reason, 1990, p. 56). Once learned, these behaviors are relatively automatic and do not rely on higher-level cognitive control nor on problem-solving processes. Errors are likely to be *slips* and *lapses* due to inattention, interference, and distraction. Among trained school psychologists, skill-based errors are likely to occur during such overlearned professional activities as administration and scoring of tests. Research has, in fact, found an alarming number of scoring errors on intelligence tests (Slate, Jones, Coulter, & Covert, 1992) as well as objective personality tests (Allard & Faust, 2000).

Rule-Based Errors

When problem solving, people learn to combine information for greater mental efficiency and to develop complex sets of if-then rules with utility in particular situations. This allows the problem solver to quickly abstract the pertinent details of a situation and automatically apply prototypical strategies that have previously been effective in similar situations. These cognitive rules can be complex. For example, expert medical diagnosticians recognize meaningful patterns that allow them to identify diseases and quickly access mental models for treatment of each disease (Ericsson, 2004).

Rules formed in this manner are not necessarily the most efficient and can go astray in a variety of ways. Initially, there are basic information processing limitations. People have a short-term memory capacity of 7 ($\pm$2) bits of information and are inaccurate when attempting to interpret the interaction of more than 3 or 4 variables (Halford, Baker, McCredden, & Bain, 2005). For example, attempts to verify the complex nonlinear configural rules claimed by clinicians have consistently found that simple linear models are equally accurate (Ruscio, 2003; Sandavol, 1998). Ultimately, the problem solver may attend to noninformative aspects of a situation, ignore or discount other important signs, persist in using a familiar but ineffectual rule, apply the wrong rule, employ rules inconsistently, and so on (McDermott, 1981). Children, for instance, often misapply rules when solving subtraction problems, revealing a systematic misunderstanding of borrowing (Reason, 1990). Likewise, a clinician may automatically, but incorrectly, diagnose learning disabilities when observing depressed Arithmetic, Coding, Information, and Digit Span subtest scores on a Wechsler scale (Watkins, 2003).

Knowledge-Based Errors

When confronted with a problem, humans prefer to search for and apply a rule-based solution. However, they revert to knowledge-based reasoning in novel situations or when available rules are not sufficient. Errors at this level arise from resource limitations and incomplete or incorrect knowledge.

Suboptimal Decisions by Psychologists

School psychologists are often called on to make difficult decisions with incomplete and uncertain data in complex environments: classification decisions, placement decisions, intervention decisions, and many others. Unfortunately, it has been well documented that school psychologists exhibit inconsistency and inaccuracy in their professional decisions (Algozzine & Ysseldyke, 1981; Aspel, Willis, & Faust, 1998; Barnett, 1988; Brown & Jackson, 1992; Davidow & Levinson, 1993; Dawes, 1994; Della Toffalo & Pedersen, 2005; deMesquita, 1992; Fagley, 1988; Fagley, Miller, & Jones, 1999; Gnys, Willis, & Faust, 1995; Huebner, 1989; Johnson, 1980; Kennedy, Willis, & Faust, 1997; Kirk & Hsieh, 2004; Macmann & Barnett, 1999; McDermott, 1980, 1981; Ysseldyke, Algozzine, Regan, & McGue, 1981).

Flawed decision making has also been found among clinical psychologists and psychiatrists. When diagnosing psychopathology, psychologists and psychiatrists have demonstrated weak interrater agreement, made errors of both under- and overidentification, and been unduly influenced by the race or gender of the client (Clark & Harrington, 1999; Faust, 1986; Garb, 1997, 1998; Garb & Boyle, 2003; Garb & Lutz, 2001). For example, "normal individuals have been misdiagnosed as brain damaged in about one out of every three cases" by neuropsychologists (Wedding & Faust, 1989, p. 241).

Psychologists have also been found to be unreliable and inaccurate in assigning children to the most appropriate level of care (Bickman, Karver, & Schut, 1997), have demonstrated low agreement when determining the function of school refusal behaviors among children (Daleiden, Chorpita, Kollins, & Drabman, 1999), and disagreed when composing case formulations for treatment of depression (Persons & Bertagnolli, 1999). Clinician agreement and accuracy on length of treatment and treatment recommendations have also been negative (Allen, Coyne, & Logure, 1990; Garb, 1998, 2005; Strauss, Chassin, & Lock, 1995).

Given this bleak record, it is not surprising that statistical prediction rules have consistently outperformed subjectively derived clinical predictions (Dawes, Faust, & Meehl, 1989; Grove & Meehl, 1996) and that experienced clinicians are no more accurate than novices in many clinical judgment tasks (Garb, 1998). In recognition of this dismal situation, Meehl (1973), in his "Why I Do Not Attend Case Conferences," scathingly satirized psychological decision making.

Suboptimal Decisions by Other Professionals

Evidence of decision-making inconstancy is not restricted to psychologists. The agreement of psychiatrists on diagnoses made in routine clinical practice and diagnoses based on semistructured interviews has been found to be poor (Shear et al., 2000). Similarly, agreement by physicians on psychotropic medication prescriptions, surgical procedures, and mammogram interpretation has been variable (Beam, Layde, & Sullivan, 1996; Dunn et al., 2005; Pappadopulos, et al., 2002). Large-scale studies of hospitalized patients have estimated that preventable medical errors annually account for 44,000 to 98,000 deaths in the United States (Leape, 1994). In agreement, autopsy studies have found high rates (35% to 40%) of erroneous diagnoses (Anderson, Hill, & Key, 1989). A meta-analysis of the reliability and validity of child protective agency caseworker decisions about allegations of child sexual abuse estimated that there are "at least 25,000 erroneous substantiation [of child sexual abuse] decisions (false positives and false negatives) per year by CPS case workers" (Herman, 2005, p. 105). Punitive monetary awards and unjust verdicts have been traced to inaccurate jury decisions (Colwell, 2005; Hastie, Schkade, & Payne, 1999). Dramatically, faulty analysis of data from space shuttle launches by NASA engineers failed to detect the o-ring malfunction caused by cold temperatures that destroyed the *Challenger* (Dawes, 2001).

Suboptimal Decisions are Universal and Systematic

In fact, research has conclusively demonstrated that *all* human decision making is susceptible to incomplete data gathering, cognitive shortcuts, errors, and biases (Arkes, 1991; Baron, 1994; Dawes, 2001; Foster & Huber, 1997; Gilovich, Griffin, & Kahneman, 2002; Nickerson, 2004; Nisbett & Ross, 1980; Plous, 1993; Reason, 1990). Whereas early conceptions of decision making were based on the fundamental belief that humans are rational, Simon (1955), who won a Nobel Prize in Economics in 1978 for his work, recognized that people do not make normatively accurate, optimal decisions because of their incomplete access to information and limited computational and predictive abilities. Instead, humans simplify the parameters of the situation, approximate the computations needed for a decision, and arrive at a satisfactory, although not necessarily optimal, decision.

Following Simon's observations, Tversky and Kahneman (1974) demonstrated that when making judgments under uncertainty people are likely to use a wide variety of nonnormative information processing techniques. These judgmental heuristics, cognitive shortcuts, or cognitive rules of thumb (Kahneman & Tversky, 1996) often yield decisions that are close approximations to optimal, but in circumstances that require logical analysis and an understanding of abstract relationships they can result in systematic biases (Baron, 1994). Kahneman's work on human judgment and decision making under uncertainty was recognized with the Nobel Prize in Economics in 2002.

Piattelli-Palmarini (1994) has referred to these heuristics and biases as *cognitive illusions* and compared them to the classical visual illusions described by experimental psychologists. For example, the vertical and horizontal lines at the top of Figure 12.1 are exactly the same length. Likewise, the two horizontal lines found at the bottom of Figure 12.1 are of equal length. Based on the limitations of human vision, a wide variety of visual illusions have been demonstrated (Robinson, 1998).

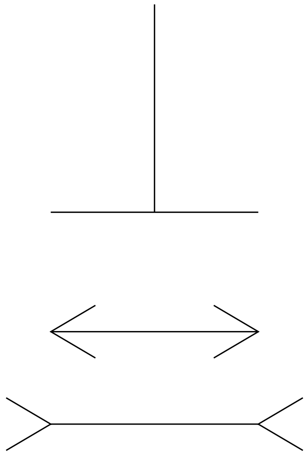

FIGURE 12.1 Visual illusions.

Similarly, a variety of cognitive illusions can be illustrated (see Plous, 1993 for multiple examples). For instance: *Each of the cards in Figure 12.2 has a number on one side and a letter on the other. Someone asserts that "if a card has a vowel on one side, it has an even number on the other side." Which of the cards should you turn over to decide whether this person is lying?* Most research participants, including psychologists, elected to look at the hidden sides of cards E and 2 in an attempt to confirm the cooccurrence of vowels and even numbers (Evans & Wason, 1976). However, the observation of an odd number on the reverse of card E or a vowel on the reverse of card 5 would more efficiently refute the statement (Plous, 1993).

A second cognitive illusion is illustrated by this diagnostic problem: *The probability of colorectal cancer is 0.3%. If a person has colorectal cancer, the probability that a hemoccult test will show a positive result is 50%. If a person does not have colorectal cancer, the probability of a positive hemoccult test is 3%. Considering only this information, what is the probability that a person who has a positive hemoccult test actually has colorectal cancer?* Many people think the correct answer is around 50%. Alarmingly, when 24 physicians were tested, only one gave the correct answer of 5% (Gigerenzer & Edwards, 2003).

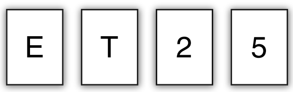

FIGURE 12.2 Given four cards, participants are asked to test the rule that, "if a card has a vowel on one side, it has an even number on the other".

Common Cognitive Heuristics, Biases, or Illusions

The perceptual illusion illustrated in Figure 12.1 persists even after using a ruler to measure the line segments. The only resolution is to use a ruler and remain confident in reason over the senses. Following this principle, a safe pilot will rely on flight instruments over fallible visuo-perceptual cues when flying in darkness. Analogous to the good pilot, the good psychologist recognizes that intuitive cognitive cues might be satisfactory in most situations but potentially misleading when making cognitively complex decisions under uncertainty. Although Croskerry (2003) identified more than 30 cognitive rules of thumb in medicine, the following cognitive heuristics, biases, confusions, and illusions are especially pertinent for school psychology.

Representativeness

According to Tversky and Kahneman (1974), people often answer probabilistic questions by relying on "the degree to which A is representative of B, that is, by the degree to which A resembles B" (p. 1124). For example, consider Steve, a man who has been described as "very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail" (Tversky & Kahneman, 1974, p. 1124). When people were asked to select Steve's probable occupation (i.e., farmer, salesman, pilot, librarian, or physician), they tended to choose librarian because the description of Steve was most *representative* of, or similar to, their stereotype of a librarian.

In diagnostic decision making, the representativeness heuristic seems to cause psychologists to discount formal diagnostic criteria in favor of a comparison of how similar the client is to the stereotypical or prototypical client with that diagnosis (Garb, 1996). Stereotypes and prototypes are partially based on clinicians' experiences, so they differ from one clinician to another and from published diagnostic criteria. Representativeness often corresponds to likelihood, so it can yield accurate results. The problem with relying on representativeness is that other relevant factors can be ignored or discounted, leading to error (Tracey & Rounds, 1999). These factors are discussed below.

Insensitivity to Prior Probabilities

The base rate, or prior probability, of a disorder or outcome has no effect on representativeness, but should have a major influence on the calculation of an accurate probability estimate. In the case of Steve, the base rate of occupations in the population can allow a more accurate prediction of whether he is, for example, a librarian or a salesman. After all, if salesmen are much more common than librarians then, absent other pertinent information, it is more rational to assign Steve to the salesman category. Interestingly, Tversky and Kahneman (1974) found that people only ignored base rates in the absence of other information. In contrast, they relied on representativeness rather than base rates when supplementary information was available. In most clinical situations, psychologists will have obtained considerable information about clients and, thus, are likely to discount or ignore base rates in favor of representativeness (Kennedy et al., 1997). Accordingly, their diagnoses tend to be subjective comparisons of the match between client symptoms and prototypical diagnostic categories (Garb, 1997). The ramifications of base rate neglect in psychodiagnostic decisions were first described by Meehl and Rosen (1955).

Misperception of Regression

In his 1877 investigation of inheritance, Sir Francis Galton found that tall parents had, on average, children who were shorter than them and short parents had, on average, children who were taller than them (Barnett, van der Pols, & Dobson, 2004). Called *regression to the mean* by Galton, this statistical phenomenon has since been found in a wide variety of situations where people are selected based upon an extreme score or characteristic. For example, a person who scored very low on one examination is likely to score somewhat higher (closer to the mean) on a second examination because extreme scores contain error that will not be repeated on a subsequent test. The problem is that "people do not develop correct intuitions about this phenomenon. First they do not expect regression in many contexts where it is bound to occur. Second, when they recognize the occurrence of regression, they often invent spurious causal explanations for it" (Tversky & Kahneman, 1974, p. 1126).

The tendency to overlook or misattribute regression effects can lead to pernicious outcomes. For example, Tversky and Kahneman (1974) described a situation where flight school instructors erroneously concluded that it was harmful to praise student pilots for outstanding performance. Nonregressive predictions are also likely to be dangerous when "measures designed to stem a 'crisis' (a sudden increase in crime, disease, or bankruptcies, or a sudden decrease in sales, rainfall, or Olympic gold medal winners) will, on the average, seem to have greater impact than there actually has been" (Nisbett & Ross, 1980, p. 163). Other detrimental outcomes of nonregressive judgments have been described by Barnett et al. (2004), Bland and Altman (1994), Glutting and McDermott (1990), and Sandoval (1998).

Misconceptions about chance

To render realistic probability estimates, an accurate understanding of how chance operates is necessary. Unfortunately, there are a variety of common misconceptions about chance.

Gambler's fallacy

People think that chance is a self-correcting process so that deviations in one direction will cause offsetting deviations in the opposite direction to restore the balance (Tversky & Kahneman, 1993). Most famously, this results in the *gambler's fallacy*, the belief that a run of bad, or good, luck will soon be followed by the opposite. Thus, after a long run of red on the roulette wheel, most people believe that black is due. Likewise, if a series of coin tosses has resulted in consecutive heads, many people expect a tail to appear on the next toss. However, each spin of the wheel or toss of the coin is independent and, thus, each outcome is also independent.

Conjunction fallacy

People tend to believe that the conjunction of two events is more, rather than less, probable than one of the events alone (Fantino, 1998). This, of course, violates the basic principle of probability (Dawes, 1993). For example, research participants were told that "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice" (Kahneman & Tversky, 1996, p. 583). Participants were then asked whether it was more likely that Linda was a (a) bank teller or (b) bank teller active in the feminist movement. Logic dictates that being both a bank teller *and* an active feminist is less likely than being either a feminist *or* a bank teller. Participants in many studies, however, chose the incorrect conjunctive response. This violates probability principles but is consistent with a bias toward representativeness.

Illusory correlation

People are not very good at distinguishing random from nonrandom outcomes and are poor judges of correlation (Baron, 1994). Studies have consistently found that people see patterns in random data and, as a consequence, have a tendency to overinterpret chance events (Gnys et al., 1995; Plous, 1993). A classic example in psychology is the *illusory correlation* phenomenon described by Chapman and Chapman (1967), who demonstrated that people associated features of projective drawings (e.g., large or unusual eyes) with diagnostic labels (e.g., suspiciousness) because of the apparent similarity, or representativeness, of the sign and symptom when, in fact, no such correlation existed. These faulty associations are strikingly similar to the shared clinical stereotypes held by many psychologists about human figure drawings and inkblots and are a prime example of how "we convince ourselves that we know all manner of stuff that just isn't so" (Paulos, 1998, p. 27).

Insensitivity to sample size

The size of the sample drawn from a population should have a major influence on estimates of the accuracy with which that sample represents the population. Small samples are likely to be variant whereas large samples are less likely to stray from population parameters. In statistics, this is called the law of *large* numbers. It appears that people believe small samples to be more representative of the population than sampling theory would suggest. For instance, when asked the probability of obtaining an average height greater than six feet in samples of 1000, 100, and 10, participants rendered the same value for all three samples (Tversky & Kahneman, 1974). This tendency to regard a sample, regardless of its size, as representative of a population has been called the law of *small* numbers (Tversky & Kahneman, 1971) and has been shown to result in exaggerated confidence in the validity of conclusions based on small samples. For example, clinicians may think a small sample of child behavior (e.g., during a three-hour testing session) generalizes to the classroom and home even though research has shown this to be an unwarranted assumption (Glutting, Youngstrom, Oakland, & Watkins, 1996). Similarly, the behavior of a small, unrepresentative group of children encountered during prior clinical experiences will be overgeneralized.

The reliance on small, idiosyncratic clinical samples can result in the *clinician's illusion* (Cohen & Cohen, 1984), where clinicians and researchers hold disparate beliefs about the long-term prognosis for a disorder. The clinician samples from the population currently suffering from the disorder and thereby obtains cases that are biased toward long duration. In contrast, research samples more nearly approximate a population sample composed of cases with all possible durations and severities.

Pseudodiagnosticity

When tasked with the estimation of the relationship between two variables based on information in 2 × 2 contingency tables (e.g., positive or negative scores on a test versus presence or absence of a disorder), people tend to place undue emphasis on the cell that represents positive test scores in the presence of a disorder and pay insufficient attention to the three other cells (Doherty, Mynatt, Tweney, & Schiavo, 1979). In the contingency table illustrated in Figure 12.3, people will disproportionately focus on the *Yes-Positive* cell. However, all four cells must be examined to accurately judge the strength of the relationship (Schustack & Sternber, 1981). Failing to do so systematically biases the estimate upward, potentially leading to the conclusion that a strong relationship exists when it does not (Nickerson, 2004).

Inverse probabilities

Insensitivity to prior probabilities and other misperceptions about chance contribute to a persistent difficulty in distinguishing conditional probabilities. For example, the probability of being a chronic smoker *conditional* on (given) a diagnosis of lung cancer is about .90, but the probability of having lung cancer *conditional* on (given) smoking is only around .10 (Dawes, 2001). Thus, many lung cancer patients are smokers but only a small minority of smokers will succumb to lung cancer. If the purpose of an analysis is to predict behavior, inverse probabilities will usually be systematic overestimates (Dawes, 1993). These results are predicated on the theorems of Thomas Bayes, an eighteenth-century British cleric (for a full explication and formulae, see Nickerson, 2004).

Within the context of medical and psychological tests, the probability of a positive test result given the presence of the disorder is known as the *sensitivity* of the test. The probability of the presence of the disorder given the positive test result is known as the *positive predictive power* of the test. Other conditional probabilities are *specificity*, which is the probability that the test is negative given that the disorder is absent, and *negative predictive power*, which is the probability that the disorder is absent given that the test is negative. These outcomes are illustrated in Figure 12.3 and described in Table 12.1. Psychologists are usually asked to predict membership in a diagnostic category given a positive test score and

TEST RESULT	DISORDER: Yes	DISORDER: No
Positive	True Positive	False Positive
Negative	False Negative	True Negative

FIGURE 12.3 Contingency table or matrix of possible outcomes when a test is used to diagnose a disorder.[1]

[1] True positives are those with the disorder who test positive. True negatives are those without the disorder who test negative. False negatives are those with the disorder but the test falsely indicates the condition is not present. False positives are those without the disorder but the test falsely indicates the condition is present.

Table 12.1 Diagnostic Statistics

Statistic	Description	Calculation[1]
Sensitivity	Given that a person has the target disorder, the probability of obtaining a positive test score.	TP ÷ (TP + FN)
Specificity	Given that a person does *not* have the target disorder, the probability of obtaining a negative test score.	TN ÷ (TN + FP)
Positive Predictive Power	Given that a person obtains a positive test score, the probability of that person having the target disorder.	TP ÷ (TP + FP)
Negative Predictive Power	Given that a person obtains a negative test score, the probability of that person *not* having the target disorder.	TN ÷ (TN + FN)
Overall Correct Classification	Hit rate. Proportion of people with *and* without the target disorder who were correctly classified by the test.	(TP + TN) ÷ (TP + TN + FP + FN)
Area Under the Curve (AUC)	Probability that a person randomly selected from the group with the disorder will have a higher score on the test than a randomly selected person from the group that does not have the disorder.	

[1] TP = true positive, FN = false negative, TN = true negative, and FP = false positive.

are rarely charged with predicting a positive test score given a particular diagnosis. Thus, the positive predictive power of the test is usually the statistic of interest, depending on the purpose of testing. Nevertheless, an understanding of the relationships between the diagnostic statistics in Table 12.1 is critical to appropriate application of tests in psychology and medicine (Galanter & Patel, 2005).

Availability

Rather than laboriously calculating the likelihood of an event or outcome from appropriate prior probabilities, people often make a judgment based on how easily they can bring examples or occurrences to mind (Tversky & Kahneman, 1974). This causes them to overestimate the probability of an easily recalled event and underestimate the probability of an ordinary or difficult to recall event. Events are more easily recalled if they are vivid, salient, visualizable, recent, imaginable, and explainable. For example, "which is a more likely cause of death in the United States—being killed by falling airplane parts or by a shark?" (Plous, 1993, p. 121). Most people think sharks are the greatest risk, whereas falling airplane parts are actually more likely. Shark attacks are easy to visualize and receive considerable media attention, which makes them easier to bring to mind. Availability also comes into play when people are faced with a choice between vivid personal testimonials versus "pallid, abstract, or statistical information" (Plous, 1993, p. 126). When considering the purchase of a new car, for example, emotional stories of a friend's lemon can outweigh the comprehensive statistical data published by *Consumer Reports*.

Research has also shown that thinking about or explaining a future event can lead to its increased availability in memory. For instance, when participants were asked to generate explanations for hypothetical future events and later judged the likelihood of those events, increased likelihood estimates for the previously explained events resulted (Hirt & Markman, 1995). Similarly, the extent to which physicians simply imagined being exposed to HIV at work subsequently increased their estimate of actual risk of exposure (Heath, Acklin, & Wiley, 1991). Likewise, illnesses particularly difficult to treat, those that received media attention, and recent conference topics were judged to be more common by physicians than they actually are (Galanter & Patel, 2005). In psychology, extreme cases from a clinician's unrepresentative sample of clients are especially memorable because they are vivid and salient, which partially explains why clinicians rely on their own experience over "pallid, abstract, [and] statistical" (Plous, 1993, p. 126) research reports (Tracey & Rounds, 1999). Likewise, preferred theories and preconceptions are readily available in memory and exert a powerful influence.

Anchoring and Adjustment

In many situations, people make estimates by starting with an initial value that is then adjusted, based on computations or further evidence, to arrive at a final solution. Unfortunately, these adjustments are often insufficient. As described by Piattelli-Palmarini (1994), "we always remain anchored to our original opinion, and we correct that view only starting from that same opinion" (p. 127). A demonstration of this bias was provided by Tversky and Kahneman (1974) who asked two groups of people to estimate the percentage of African countries in the United Nations. Before that, however, both groups had spun a roulette wheel to obtain a random comparison number. The first group randomly began with 10 and the second group with 65. Their subsequent estimates of the percentage of African countries in the UN were 25 and 45, respectively. These results have been replicated in more realistic situations. For example, real estate agents were shown to be anchored to initial listing prices (Northcraft & Neale, 1987). As summarized by Plous (1993, p. 151), "the effects of anchoring are pervasive and extremely robust.... People adjust insufficiently from anchor values, regardless of whether the judgment concerns the chances of nuclear war, the value of a house, or any number of other topics."

The adjustment and anchoring heuristics may become influential when school psychologists are provided with initial information about referrals. For example, "referral information, or previous test data may be available that may unduly affect the ultimate outcome of an assessment by providing a different 'starting point' in a way analogous to the anchoring of numerical predictors" (Fagley, 1988, p. 317). This supposition has been supported by several studies (deMesquita, 1992; Della Toffalo & Pedersen, 2005; McCoy, 1976; Ysseldyke et al., 1981).

Framing

Scores of studies have demonstrated that people view positive outcomes as more probable than negative outcomes (Plous, 1993). For example, Rosenhan and Messick (1966) asked participants to predict the probability of drawing cards from a deck containing cards stamped with smiling or frowning faces. People consistently underestimated the probability of obtaining a card with a frowning face and overestimated the probability of drawing a card with a smiling face. Likewise, people unrealistically expect their personal outcomes to exceed those of other people or objective indicators (Carroll, Sweeny, & Shepperd, 2006). For instance, students rated themselves more likely than others to experience positive life events and less likely to experience negative life events (Weinstein, 1980), smokers overestimated the probability that they would quit smoking in the coming year (Weinstein, Slovic, & Gibson, 2004), and student teachers believed they would experience less difficulty than the average beginning teacher during their first year of teaching (Weinstein, 1988). This *optimistic* bias might be salient when school psychologists consider the advisability of intervention or placement options.

However, framing effects are more complex than a simple preference for positive outcomes. Tversky and Kahneman (1981) argued that people evaluate outcomes in context: positive framing (describing options in terms of gains) leads people to be risk averse and choose certain gain over potential loss, whereas negative framing (describing options in terms of losses) causes people to accept risk to avoid certain loss. Thus, different ways of presenting the same information can result in diametrically different decisions. In one study, for example, participants were presented two sets of data and asked to make two separate choices. First, would they prefer (a) a sure gain of $75 or (b) a 75% chance to win $100 and a 25% chance to gain nothing. Second, would they prefer (c) a sure loss of $75 or (d) a 75% chance to lose $100 and a 25% chance to lose nothing. In the first choice, 84% of the participants chose the first alternative (a sure gain). In contrast, 88% of the participants chose to gamble against a sure loss in the second scenario. As expected, people were risk averse when gains were at stake, but chose risk when losses were at issue.

It has been discovered that the way medical procedures are framed (mortality versus survival) has a powerful effect on patient and physician choices (Armstrong, Schwartz, Fitzgerald, Putt, & Ubel, 2002), and the way bets are framed (winning versus losing) profoundly affects gamblers (Nickerson, 2004). Framing effects have also been observed in school psychology. Fagley et al. (1999) provided doctoral school psychology students with positively and negatively framed choices on dropout prevention programs, smoking prevention programs, mainstreaming versus special education placements, and so on. As expected, more risky choices were made in response to negatively framed decision problems.

Hindsight Bias

Probability estimation can also be distorted by hindsight bias (Fischhoff, 1975), where people who know the outcome of an event will posthoc overestimate the probability of that outcome. This is analogous to Monday morning quarterbacking, where many people are confident that they could have predicted the outcome of Sunday's football game. However, this is an overestimate of their actual prognostication abilities. In part, this *illusion of learning* stems from an inability to imagine an alternative outcome (i.e., availability heuristic). Unfortunately, people seem to be relatively insensitive to the operation of hindsight bias and it may cause them to become more confident of their decisions because the actual outcomes seemed so obvious and preordained. For example, hindsight knowledge of a diagnosis might result in an overinflated belief that one would have been able to make the diagnosis with accuracy (Wedding & Faust, 1989). Hindsight bias could also cause clinicians to remember successful predictions of client behavior and forget or ignore unsuccessful predictions (Gibbs & Gambrill, 1996).

Fundamental Attribution Error

There is a robust, ubiquitous tendency of people to (a) attribute the behavior of others to enduring and consistent dispositions (i.e., personal traits) rather than the particular situation, and (b) attribute their own behavior to the demands of the situation instead of personal traits (Kahneman & Tversky, 1996). Thus, people will attribute success to their own efforts, intelligence, and perspicacity and failure to bad luck and circumstances beyond their control. For instance, one study found that 97.3% of student problems were attributed by teachers to internal student traits and home causes (Christenson, Ysseldyke, Wang, & Algozzine, 1983). Another study found that all student learning problems were attributed by school psychologists to student or family deficiencies and none to school deficits (Alessi, 1988). In clinical practice, the fundamental attribution error "results in blaming the client, rather than identifying and altering environmental events related to problems" (Gibbs & Gambrill, 1996, p. 132).

Overconfidence

Possibly because of fundamental attributions and other cognitive biases, people often express extreme confidence in highly fallible judgments (Smith & Dumont, 2002). In fact, "judgments produced in decision environments such as psychodiagnosis, which are by nature complex and ambiguous, appear to be most vulnerable to overconfidence" (Smith & Dumont, 1997, p. 342). Surprisingly, research has shown that confidence is directly related to the *number* of decisions made, irrespective of the *accuracy* of those decisions (Arkes, Hackett, & Boehm, 1989). Thus, professionals may become more confident with experience but not more accurate (Dunning, Heath, & Suls, 2004). Because of this overly positive view of themselves and their decisions, individuals tend to think their own behavior is typical of others and to assume that most people would have made the same decision as themselves (the *false consensus effect*). Overconfidence can impart a false sense of security and overconfident individuals may be less likely to objectively evaluate their own performance. Physicians, nurses, police officers, and psychologists have been shown to exhibit unwarranted confidence in their professional decisions (Baumann, Deber, & Thompson, 1991; Garb & Schramke, 1996; Kassin & Gudjonsson, 2004).

Confirmation Bias

It has repeatedly been found that people prefer confirmatory to disconfirmatory strategies and, accordingly, selectively seek and interpret evidence supportive of their prior beliefs or hypotheses and ignore or discount nonsupportive evidence (Nickerson, 1998). Typically, a preliminary hypothesis is quickly formed and then support for that initial position becomes the salient activity. As described by Baron (1994, p. 302), "this is what makes us into lawyers, hired by our own earlier views to defend them against all accusations, rather than detectives seeking the truth itself."

Confirmation bias does not operate alone—it is confounded with other cognitive heuristics, biases, and misperceptions. Representativeness, hindsight bias, fundamental attributions, and availability, among other phenomena, operate in the formulation of initial hypotheses. Research has demonstrated that counselors, jurors, lawyers, physicians, police officers, and psychologists develop preliminary hypotheses very quickly and inadequately revise them based on subsequent information (Colwell, 2005; Haverkamp, 1993; Kassin & Gudjonsson, 2004; Lopez, 1989; Meehl, 1960). For example, psychologists arrived at problem formulations within the first few clinical sessions, and failed to reformulate them during the subsequent 24 sessions (Meehl, 1960). This combination of hastily formed diagnoses and resistance to reasonable competing alternatives was the most common cognitive cause of diagnostic errors made by internists (Graber, Franklin, & Gordon, 2005).

Resistance to revision of initial hypotheses arises from several sources. First, people seek the information they expect to find, assuming that their hypotheses are correct. Accordingly, they will disproportionately emphasize the *Yes-Positive* cell (pseudodiagnosticity) of Figure 12.3 and fail to fully consider the diagnostic information contained in the other three cells. They will also favor positive tests, as illustrated in Figure 12.2. Third, they will ignore base rates, misinterpret regression effects, overgeneralize from small samples of behavior or clients, find explanations for chance occurrences, frame the problem in positive terms, and so forth. Fourth, information acquired early in the decision-making process will be given more weight than information acquired later. This *primacy effect* will prematurely terminate the revision process. Finally, even if initial hypotheses are revised, those adjustments will be insufficient because they were anchored to early, inaccurate estimates.

Once expectations have been confirmed, individuals become increasingly overconfident. They consistently focus on evidence that affirms their accuracy and disregard contrary data. Because of this heightened confidence, a false sense of security ensues. In this flattering light, individuals assume that most people would have made the same decision as them and consequently see no need to objectively evaluate their own performance. Through these intertwined mechanisms, people develop high levels of confidence in their decisions regardless of the objective merit of those decisions.

INTERVENTIONS

Although the variety and scope of human errors have been extensively investigated, interventions to decrease errors have not received commensurate attention. Few interventions have been investigated and even fewer have been found effective. Within that context, the following recommendations are proffered to improve diagnostic decision making and clinical judgment in school psychology.

Understand Cognitive Heuristics

School psychologists must become familiar with cognitive biases and heuristics on the assumption that such awareness will reduce the influence of those cognitive biases and heuristics. Although not sufficient, knowledge is necessary. The comprehensive treatments of Nickerson (2004) and Plous (1993), as well as direct instruction via classes and continuing education programs may be informative (Croskerry, 2003; Davidow & Levinson, 1993; Lilienfeld, Lynn, & Lohr, 2003).

Check for Errors

School psychologists must also recognize their vulnerability to skill-based errors, especially when administering and scoring tests. Such errors are indefensible. Scoring and administration errors can be ameliorated by careful use of checklists and guidelines that provide immediate corrective feedback (Moon, Fantuzzo, & Gorsuch, 1986). Computer-based scoring may also be beneficially employed.

Acknowledge Limitations

School psychologists must acknowledge that they are prey to the same sensory and cognitive limitations as other professionals. Meehl (1973) wrote, "it is absurd, as well as arrogant, to pretend that acquiring a PhD somehow minimizes me from the errors of sampling, perception, recording, retention, retrieval, and inference to which the human mind is suspect" (p. 278). Even the visual judgment of graphed data in single-case research designs by expert behavioral analysts has been found to be unreliable (Kromrey & Foster-Johnson, 1996). Understanding these limitations will be especially important as school psychologists increasingly utilize single-case design methodology for decision making within response-to-intervention approaches promulgated by special education laws. In terms of information processing, this means that configural complexity, limits of memory, and so on are especially relevant. Therefore, psychologists should use checklists, flowcharts, notes, practice guidelines, computer programs, standardized work processes, and other aids to reduce reliance on memory and minimize the influence of other cognitive biases (Galanter & Patel, 2005).

Become a Self-Directed, Lifelong Learner

School psychologists must become "informed students of the professional literature, capable of assessing, understanding, and applying the quality of evidence therein" (Gill & Pratt, 2005, p. 95). This includes knowledge of assessment and intervention practices, as well as of measurement and statistical principles. Good intentions, without knowledge, do not assure beneficial results. Quite the opposite, harm can occur: iatrogenic effects have been observed with well-intentioned crisis intervention and adolescent conduct disorder programs (Bootzin & Bailey, 2005; Gambrill, 2005). Professional practice extends for decades after initial preparation, and the half life of scientific knowledge is steadily decreasing (Rutter & Yule, 2002). One study estimated that college exposes professionals to only about one-sixth of the knowledge they will need during their careers (Tenopir & King, 1997). Consequently, it is imperative that school psychologists manage their own learning. The increasing availability of electronic documents and search engines may be useful tools for such self-directed, lifelong learning.

Avoid Overconfidence

Although comforting, overconfidence in complex professional judgments is unwarranted. Research has taught us that we are likely to attribute success to our own astuteness, but failure to bad luck and environmental circumstances. In contrast, we are likely to attribute the behavior of others to enduring dispositions. Consequently, *our* decisions will appear ineluctably reasonable to *us*. This is especially true in hindsight, where it can seem as if the decision was so obvious that most other clinicians would have inevitably arrived at the same conclusion. Ironically, the very processes that generate overconfidence may operate to shield a person from recognizing the limits of their competence. As noted by Kruger and Dunning (1999), the deficits in metacognitive skills that allowed an *illusion of validity* to develop are implicated in the inability to recognize a discrepancy between behavior and belief. However, appropriate feedback and training appear to diminish overconfidence (Baumann et al., 1991; Smith & Dumont, 1997).

Experience is not Necessarily Expertise

Unfortunately, the overconfident perception of competence can be exacerbated by experience. "More experienced clinicians are more confident of their judgments than are novices, even though the judgments are no less accurate" (Tracey & Rounds, 1999, p. 125). Although experience can be valuable, it generally does not produce expertise unless acquired under specific conditions. Most importantly, expert performance only develops after about 10 years of deliberate practice with feedback (Ericsson & Charness, 1994). This has been seen in diverse fields, including chess, music, medicine, and athletics (Ericsson, 2004). For example, chess grandmasters had studied about five times more than average chess tournament players by their tenth year of play (Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005). Without deliberate practice, 10 years of experience "may lead to nothing more than learning to make the same mistakes with increasing confidence" (Skrabanek & McCormick, 1990, p. 28) or may have no more value than one year of experience, repeated ten times. This may be one explanation for the ineffectiveness of typical clinical services (Bickman, 1999).

Consequently, school psychologists must not assume, absent empirical evidence, that their experience equates to competence (Dawes, 1994; Garb & Boyle, 2003). Clinicians cannot achieve expertise without extensive deliberate practice. Deliberate practice requires immediate corrective feedback. Clinicians do not usually receive adequate feedback about their decisions and, consequently, have trouble learning from their experiences (Bickman, 1999). Although constrained by the clinical arena, psychologists must collect data on the accuracy of their decisions (Stricker, 2006) and must learn from their mistakes (Popper, 1992). Proactive attention to evaluative data is crucial to the development of professional expertise. Objective competence is infinitely preferable to overconfidence and wishful thinking.

Use Decision Aids and Actuarial Methods

McDermott (1981) found that school psychologists' decisions were affected by inconsistent decision rules, theoretical orientations, weighing of diagnostic cues, and diagnostic styles. These inconsistencies arise, at least partially, from the fallibility of human information processing. Decision aids such as checklists will help reduce inconsistency in some situations. Another powerful remedy can be found in actuarial methods, which have consistently been found superior to clinical judgment (Dawes et al., 1989). Actuarial methods depend on mechanical or statistical prediction whereas clinical methods rely on subjective, impressionistic judgment. Critically, the information used for prediction, whether actuarial or clinical, can be of any type. The task is, "given a data set (e.g., life history facts, interview ratings, ability test scores, MMPI profiles, nurses' notes), how is one to put these various facts (or first-order inferences) together to arrive at a prediction about the individual" (Grove & Meehl, 1996, p. 299). It is not the nature of the data, but how the data are combined, mechanically or clinically, that produces superior results (Baron, 1994). For example, case managers' clinical judgment regarding home treatment visits needed by aggressive and disruptive children was inferior to a linear combination of six rating items assessing parental functioning rendered by those same clinicians (Bierman, Nix, Murphy, & Maples, 2006). Accordingly, spurious arguments against actuarial methods should be rejected (Grove & Meehl, 1996) and actuarial methods applied whenever possible.

Use Reliable, Valid Tools

It is a truism that the ability of school psychologists is limited by the reliability and validity of the tools they employ (Palmiter, 2004). As illustrated by Frazier and Youngstrom (2006), reliance on instruments with solid reliability and validity evidence is integral to evidence-based diagnosis and evidence-based testing (Mash & Hunsley, 2005; McFall, 2005). School psychologists might profitably emulate this approach. Further, given that judgments are inordinately influenced by information obtained early in the diagnostic process, it is advisable to begin that process with the most valid three or four nonredundant pieces of data (Wedding & Faust, 1989).

Use Bayesian Reasoning

Bayesian reasoning entails being aware of base rates as well as avoiding inverse probabilities and pseudodiagnosticity. Unfortunately, people do not intuitively grasp Bayesian methods and have considerable difficulty applying them correctly even after professional training (Gigerenzer, 2002). For example, problems similar to the previously presented colorectal cancer example were solved by only 5% to 17% of physicians (Gigerenzer, 2002). Although people tend to do better with natural frequencies than with probabilities or percentages, decision aids such as computer programs are probably the best solution to this natural human weakness. Software solutions for calculating Bayesian statistics are freely available at www.public.asu.edu/~mwwatkin.

Do Not Confuse Classical Validity with Diagnostic Utility

Relatedly, school psychologists should not confuse classical validity methods with diagnostic statistics (Wiggins, 1988). Average group score differences indicate that *groups* can be discriminated. This classical validity approach cannot be uncritically extended to conclude that mean group differences are distinctive enough to differentiate among *individuals*. Figure 12.4 illustrates this dilemma. It displays hypothetical score distributions of children from regular and exceptional student populations. Group mean differences are clearly discernable, but the overlap between distributions makes it difficult to accurately identify group membership for those individuals within the overlapping distributions. Group separation is necessary but not sufficient for accurate decisions about individuals.

Unfortunately, errors in assigning individuals to normal or disabled groups are unavoidable given the imperfect tools and taxonomies available to psychologists (Zarin & Earls, 1993). The relative proportion of correct and incorrect diagnostic decisions depends on the cut score used. In Figure 12.4, for example, X1 represents a low cut score and X2 a high cut score. With a low cut score, there are a large number of false positive and a small number of false negative decisions. With a high cut score, there are a large number of false negative and a small number of false positive decisions. Beyond cut scores, the accuracy of diagnostic decisions is dependent on the base rate or prevalence of the particular disability in the population being assessed (Meehl & Rosen, 1955).

In contrast, by systematically using all possible cut scores of a diagnostic test and graphing true positive against false positive decision rates for each cut score, the full range of that test's diagnostic utility can be displayed (McFall & Treat, 1999; Swets, Dawes, & Monahan, 2000). Designated the receiver operating characteristic

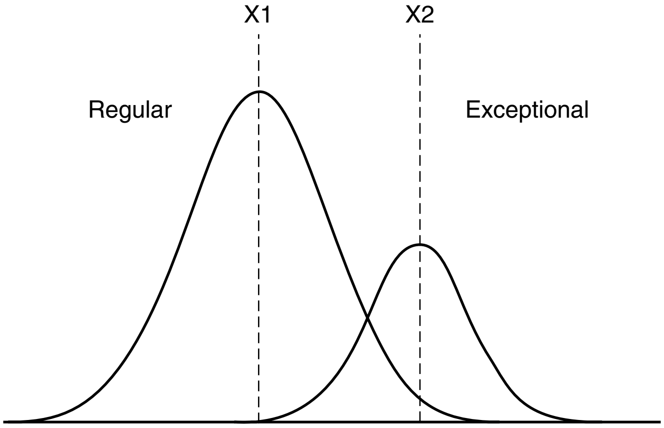

FIGURE 12.4 Hypothetical test score distributions for regular and exceptional student populations with two cut points (X1 and X2).

(ROC), this procedure is not confounded by cut scores or prevalence rates. Consequently, ROC curves are "the state-of-the-art method for describing the diagnostic accuracy of a test" (Weinstein, Obuchowski, & Lieber, 2005, p. 16) and are "recognized widely as the most meaningful approach to quantify the accuracy of diagnostic information and diagnostic decisions" (Metz & Pan, 1999, p. 1).

The area under the ROC curve (AUC) provides an accuracy index of the test. AUC values can range from 0.5 to 1.0. An AUC value of 0.5 signifies that no discrimination exists. In this case, the ROC curve lies on the main diagonal of the graph and the diagnostic system is functioning at the level of chance. In contrast, an AUC value of 1.0 denotes perfect discrimination. The AUC also has an intuitive meaning: If one person is randomly selected from the nondisordered population and one from the disordered population, the AUC is the probability of distinguishing between those two individuals with the test.

Two illustrative ROC curves are presented in Figures 12.5 and 12.6. In Figure 12.5, Verbal and Performance IQ score differences on the Wechsler Intelligence Scale for Children–Third Edition (WISC-III; Wechsler, 1991) were compared between 1,153 children with learning disabilities and the 2,200 children in the WISC-III normative sample. Figure 12.6 displays the ROC curve for the Overreactivity Scale of the Adjustment Scales for Children and Adolescents (ASCA; McDermott, Marston, & Stott, 1993) for 21 children with emotional disabilities compared to 1,056 children in the ASCA normative sample. The AUCs were .57 and .94, respectively, for Figures 12.5 and 12.6. AUC values of 0.5 to 0.7 indicate low test accuracy, 0.7 to 0.9 indicate moderate test accuracy, and 0.9 to 1.0 indicate high test accuracy (Swets, 1988). In this example, Verbal-Performance IQ score differences were not useful in identifying children with learning disabilities but ASCA Overreactivity scores were extremely accurate in distinguishing children with emotional disabilities. Accordingly, school psychologists should routinely use diagnostic statistics, including the ROC and its AUC, when considering the accuracy of diagnostic information. Software solutions for ROC curves are freely available at www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html and www.public.asu.edu/~mwwatkin.

Rely on a Scientific Approach

Adopt and adhere to a scientific approach when making professional decisions (Dawes, 1995; Hayes, Barlow, & Nelson-Gray, 1999; McFall, 1991, 2000). Unfortunately, scientific reasoning does not come naturally: it demands critical thinking, tolerance of ambiguity, skepticism, openness to criticism, and acceptance of fallibility (Baron, 1994; Cromer, 1993; Wilson, 1995). Science does not confuse reasoning with rationalizing, nor beliefs with facts. Instead, science is a self-correcting process of objective investigation and logical inquiry used to accumulate a reliable body of knowledge (Gibbs & Gambrill, 1996).

It is often assumed that hypothetico-deductive reasoning in psychological practice is analogous to scientific reasoning. That is, formulation of diagnostic hypotheses that guide

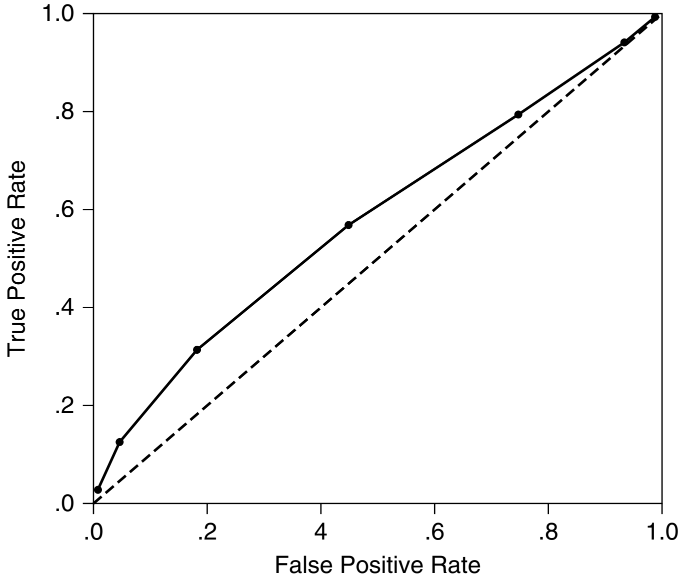

FIGURE 12.5 ROC curve of Verbal-Performance IQ score differences on the Wechsler Intelligence Scale for Children–Third Edition (WISC-III) for 1,153 children with learning disabilities and the 2,200 children in the WISC-III normative sample.

subsequent data gathering that, in turn, either supports or fails to support the proposed hypotheses. For example, Lichtenberger (2006, p. 27) suggested that "a hypothesis . . . can be confirmed with one piece of supplementary data, but two pieces of confirmatory data are preferable. If one or more pieces of data contradict the hypothesis, then that hypothesis may not be valid for that client." Hypothetico-deductive strategies work in science because they are public and can be tested and refuted by other researchers but "the idealized process often goes astray" (Aspel et al.,

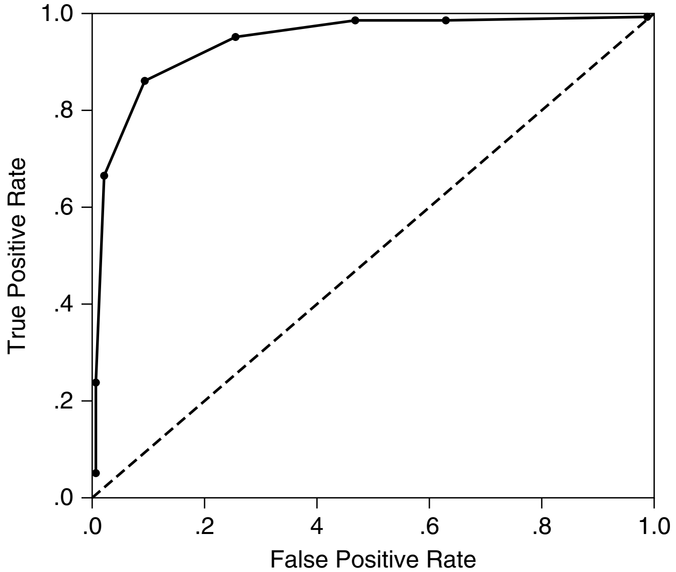

FIGURE 12.6 ROC curve of the Overreactivity Scale of the Adjustment Scales for Children and Adolescents (ASCA) for 21 children with emotional disabilities compared to 1,056 children in the ASCA normative sample.

1998, p. 138) in clinical practice. Clinicians do not apply the methodological controls against bias that are present in research settings. Further, clinical hypotheses are private (no competitors will attempt to refute the psychologist's hypotheses) and are, therefore, exceptionally vulnerable to confirmation bias.

In contrast, the criterion of falsifiability is a hallmark of scientific inquiry (Platt, 1964). That is, theories are scientific only if they can be subjected to tests that can refute them (Gibbs & Gambrill, 1996). Following this principle, clinicians must actively search for *disconfirmatory* evidence to reduce the influence of confirmatory bias (Arkes, 1991; Faust, 1986; Sandoval, 1998). "Always look first for that which disconfirms your beliefs; then look for that which supports them. Look with equal diligence for both. Doing so will make the difference between scientific honesty and artfully supported propaganda" (Gibbs, 2003, p. 89). An even better strategy may be to formulate plausible competing hypotheses and actively search for disconfirmatory evidence (Croskerry, 2003; Hirt & Markman, 1995; Plous, 1993; Tracey & Rounds, 1999; Wedding & Faust, 1989). This *competing hypothesis strategy* is preferable to the hypothetico-deductive strategy in clinical practice.

Reliance on science (both as a problem-solving process and as a reliable body of knowledge) is central to evidence-based practice, empirically supported practice, and science-based psychology (Gambrill, 2005; Gibbs & Gambrill, 1996; Lilienfeld & O'Donohue, 2006; Lonigan, Elbert, & Johnson, 1998; McFall, 1991; Stricker, 2006; USDOE, 2003). As with evidence-based assessment, school psychologists might beneficially incorporate these approaches into their professional practice.

Conclusion

Almost 300 years ago, Alexander Pope reminded us that to err is human. Nevertheless, it is the moral, ethical, and legal obligation of psychologists to err as little as possible in their diagnostic decision making and clinical judgments (Gambrill, 2005; Hummel, 1999; McFall, 2000; Meehl, 1973; Poortinga & Soudijn, 2003; Popper, 1992). Lawyers and judges are increasingly aware of scientific method and are being trained to adjudicate complex scientific disputes (Federal Judicial Center, 2000; Foster & Huber, 1997). Consequently, if psychologists ignore their duty to minimize professional error, then "some smart lawyers and sophisticated judges will either discipline or discredit us" (Meehl, 1997, p. 98).

References

Alessi, G. (1988). Diagnosis diagnosed: A systemic reaction. *Professional School Psychology*, *3*, 145–151.

Algozzine, B., & Ysseldyke, J. E. (1981). Special education services for normal children: Better safe than sorry? *Exceptional Children*, *48*, 238–243.

Allard, G., & Faust, D. (2000). Errors in scoring objective personality tests. *Assessment*, *7*, 119–129.

Allen, J. G., Coyne, L., & Logure, A. M. (1990). Do clinicians agree about who needs extended psychiatric hospitalization? *Comprehensive Psychiatry*, *31*, 355–362.

Anderson, R. E., Hill, R. B., & Key, C. R. (1989). The sensitivity and specificity of clinical diagnostics during five decades: Toward an understanding of necessary fallibility. *Journal of the American Medical Association*, *261*, 1610–1617.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin*, *110*, 486–498.

Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, *2*, 81–94.

Armstrong, K., Schwartz, J. S., Fitzgerald, G., Putt, M., & Ubel, P. A. (2002). Effect of framing as gain versus loss on understanding and hypothetical treatment choices: Survival and mortality curves. *Medical Decision Making*, *22*, 76–83.

Aspel, A. D., Willis, W. G., & Faust, D. (1998). School psychologists' diagnostic decision-making processes: Objective-subjective discrepancies. *Journal of School Psychology*, *36*, 137–149.

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2004). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, *34*, 215–220.

Barnett, D. W. (1988). Professional judgment: A critical appraisal. *School Psychology Review*, *17*, 658–672.

Baron, J. (1994). *Thinking and deciding* (2nd ed.). NY: Cambridge University Press.

Baumann, A. O., Deber, R. B., & Thompson, G. G. (1991). Overconfidence among physicians and nurses: The 'micro-certainty, macro-uncertainty' phenomenon. *Social Science in Medicine*, *32*, 167–174.

Beam, C. A., Layde, P. M., & Sullivan, D. C. (1996). Variability in the interpretation of screening mammograms by U.S. radiologists: Findings from

a national sample. *Archives of Internal Medicine*, *156*, 209–213.
Bickman, L. (1999). Practice makes perfect and other myths about mental health services. *American Psychologist*, *54*, 965–978.
Bickman, L., Karver, M. S., & Schut, J. A. (1997). Clinician reliability and accuracy in judging appropriate level of care. *Journal of Consulting and Clinical Psychology*, *65*, 515–520.
Bierman, K. L., Nix, R. L., Murphy, S. A., & Maples, J. J. (2006). Examining clinical judgment in an adaptive intervention design: The Fast Track program. *Journal of Consulting and Clinical Psychology*, *74*, 468–481.
Bland, J. M., & Altman, D. G. (1994). Some examples of regression towards the mean. *British Medical Journal*, *309*, 780.
Bootzin, R. R., & Bailey, E. T. (2005). Understanding placebo, nocebo, and iatrogenic treatment effects. *Journal of Clinical Psychology*, *61*, 871–880.
Brown, R. T., & Jackson, L. A. (1992). Ex-Huming an old issue. *Journal of School Psychology*, *30*, 215–221.
Carroll, P., Sweeny, K., & Shepperd, J. A. (2006). Forsaking optimism. *Review of General Psychology*, *10*, 56–73.
Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, *72*, 193–204.
Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology*, *19*, 151–165.
Christenson, S., Ysseldyke, J. E., Wang, J. J., & Algozzine, B. (1983). Teachers' attributions for problems that result in referral for psychoeducational evaluation. *Journal of Educational Research*, *76*, 174–180.
Clark, A., & Harrington, R. (1999). On diagnosing rare disorders rarely: Appropriate use of screening instruments. *Journal of Child Psychology and Psychiatry*, *40*, 287–290.
Cohen, P., & Cohen, J. (1984). The clinician's illusion. *Archives of General Psychiatry*, *41*, 1178–1182.
Colwell, L. H. (2005). Cognitive heuristics in the context of legal decision making. *American Journal of Forensic Psychology*, *23*(2), 17–41.
Cromer, A. (1993). *Uncommon sense: The heretical nature of science*. NY: Oxford University Press.
Croskerry, P. (2003). The importance of cognitive errors in diagnosis and strategies to minimize them. *Academic Medicine*, *78*, 775–780.
Daleiden, E. L., Chorpita, B. F., Kollins, S. H., & Drabman, R. S. (1999). Factors affecting the reliability of clinical judgments about the function of children's school-refusal behavior. *Journal of Clinical Child Psychology*, *28*, 396–406.
Davidow, J., & Levinson, E. M. (1993). Heuristic principles and cognitive bias in decision making: Implications for assessment in school psychology. *Psychology in the Schools*, *30*, 351–361.
Dawes, R. M. (1993). Prediction of the future versus an understanding of the past: A basic asymmetry. *American Journal of Psychology*, *106*, 1–24.
Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. NY: Free Press.
Dawes, R. M. (1995). Standards of practice. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 31–43). Reno, NV: Context Press.
Dawes, R. M. (2001). *Everyday irrationality: How pseudo-scientists, lunatics, and the rest of us systematically fail to think rationally*. Boulder, CO: Westview Press.
Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674.
Della Toffalo, D. A., & Pedersen, J. A. (2005). The effect of a psychiatric diagnosis on school psychologists' special education eligibility decisions regarding emotional disturbance. *Journal of Emotional and Behavioral Disorders*, *13*, 53–60.
deMesquita, P. (1992). Diagnostic problem solving of school psychologists: Scientific method or guesswork? *Journal of School Psychology*, *30*, 269–291.
Doherty, M. E., Mynatt, C. R., Tweney, R. D., & Schiavo, M. D. (1979). Pseudodiagnosticity. *Acta Psychologica*, *43*, 111–121.
Dunn, W. R., Schackman, B. R., Walsh, C., Lyman, S., Jones, E. C., Warren, R. F., & Marx, R. G. (2005). Variation in orthopaedic surgeons' perceptions about the indications for rotator cuff surgery. *Journal of Bone and Joint Surgery*, *87*, 1978–1984.
Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*, 69–106.
Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(10), S70–S81.
Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725–747.
Evans, J., & Wason, P. C. (1976). Rationalization in a reasoning task. *British Journal of Psychology*, *67*, 486–497.
Fagley, N. S. (1988). Judgmental heuristics: Implications for the decision making of school psychologists. *School Psychology Review*, *17*, 311–321.

Fagley, N. S., Miller, P. M., & Jones, R. N. (1999). The effect of positive or negative frame on the choices of students in school psychology and educational administration. *School Psychology Quarterly*, *14*, 148–162.

Fantino, E. (1998). Judgment and decision making: Behavioral approaches. *The Behavior Analyst*, *21*, 203–218.

Faust, D. (1986). Research on human judgment and its application to clinical practice. *Professional Psychology: Research and Practice*, *17*, 420–430.

Federal Judicial Center. (2000). *Reference manual on scientific evidence* (2nd ed.). Washington, DC: Author.

Fischhoff, B. (1975). Hindsight ≠ foresight: The effects of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

Foster, K. R., & Huber, P. W. (1997). *Judging science: Scientific knowledge and the federal courts*. Cambridge, MA: MIT Press.

Frazier, T. W., & Youngstrom, E. A. (2006). Evidence-based assessment of attention-deficit/hyperactivity disorder: Using multiple sources of information. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 614–620.

Galanter, C. A., & Patel, V. L. (2005). Medical decision making: A selective review for child psychiatrists and psychologists. *Journal of Child Psychology and Psychiatry*, *46*, 675–689.

Gambrill, E. (2005). *Critical thinking in clinical practice: Improving the quality of judgments and decisions* (2nd ed.). Hoboken, NJ: Wiley.

Garb, H. N. (1996). The representativeness and past-behaviors heuristics in clinical judgment. *Professional Psychology: Research and Practice*, *27*, 272–277.

Garb, H. N. (1997). Race bias, social class bias, and gender bias in clinical judgment. *Clinical Psychology: Science and Practice*, *4*, 99–120.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: American Psychological Association.

Garb, H. N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology*, *1*, 67–89.

Garb, H. N., & Boyle, P. A. (2003). Understanding why some clinicians use pseudoscientific methods: Findings from research on clinical judgment. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 17–38). New York: Guilford.

Garb, H. N., & Lutz, C. (2001). Cognitive complexity and the validity of clinicians' judgments. *Assessment*, *8*, 111–115.

Garb, H. N., & Schramke, C. J. (1996). Judgment research and neuropsychological assessment: A narrative review and meta-analyses. *Psychological Bulletin*, *120*, 140–153.

Gibbs, L. E. (2003). *Evidence-based practice for the helping professions: A practical guide with integrated multimedia*. Pacific Grove, CA: Brooks/Cole.

Gibbs, L., & Gambrill, E. (1996). *Critical thinking for social workers: Exercises for the helping professions*. Thousand Oaks, CA: Pine Forge Press.

Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you*. New York: Simon & Schuster.

Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: From innumeracy to insight. *British Medical Journal*, *327*, 741–744.

Gill, K. J., & Pratt, C. W. (2005). Clinical decision making and the evidence-based practitioner. In R. E. Drake, M. R. Merrens, & D. W. Lynde (Eds.), *Evidence-based mental health practice: A textbook* (pp. 95–122). New York: W. W. Norton.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.

Glutting, J. J., & McDermott, P. A. (1990). Principles and problems in learning potential. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment: 1. Intelligence and achievement* (pp. 296–347). New York: Guilford Press.

Glutting, J. J., Youngstrom, E. A., Oakland, T., & Watkins, M. W. (1996). Situational specificity and generality of test behaviors for samples of normal and referred children. *School Psychology Review*, *25*, 94–107.

Gnys, J. A., Willis, W. G., & Faust, D. (1995). School psychologists' diagnoses of learning disabilities: A study of illusory correlation. *Journal of School Psychology*, *33*, 59–73.

Graber, M. L., Franklin, N., & Gordon, R. (2005). Diagnostic error in internal medicine. *Archives of Internal Medicine*, *165*, 1493–1499.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law*, *2*, 293–323.

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science*, *16*, 70–76.

Hastie, R., Schkade, D. A., & Payne, J. W. (1999). Juror judgments in civil cases: Hindsight effects on judgments of liability for punitive damages. *Law and Human Behavior*, *23*, 597–614.

Haverkamp, B. E. (1993). Confirmatory bias in hypothesis testing for client-identified and counselor identified self-generated hypotheses. *Journal of Counseling Psychology*, *40*, 303–315.

Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). *The scientist-practitioner: Research and accountability in the age of managed care* (2nd ed.). Boston: Allyn and Bacon.

Heath, L., Acklin, M., & Wiley, K. (1991). Cognitive heuristics and AIDS risk assessment among physicians. *Journal of Applied Social Psychology*, *21*, 1859–1867.

Herman, S. (2005). Improving decision making in forensic child sexual abuse evaluations. *Law and Human Behavior*, *29*, 87–120.

Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, *69*, 1069–1086.

Huebner, E. S. (1989). Errors in decision-making: A comparison of school psychologists' interpretations of grade equivalents, percentiles, and deviation IQs. *School Psychology Review*, *18*, 51–55.

Hummel, T. J. (1999). The usefulness of tests in clinical decisions. In J. W. Lichtenberg & R. K. Goodyear (Eds.), *Scientist-practitioner perspectives on test interpretation* (pp. 59–112). Boston: Allyn and Bacon.

Johnson, V. M. (1980). Analysis of factors influencing special educational placement decisions. *Journal of School Psychology*, *18*, 191–202.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582–591.

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, *5*, 33–67.

Kennedy, M. L., Willis, W. G., & Faust, D. (1997). The base-rate fallacy in school psychology. *Journal of Psychoeducational Assessment*, *15*, 292–307.

Kirk, S. A., & Hsieh, D. K. (2004). Diagnostic consistency in assessing conduct disorder: An experiment on the effect of social context. *American Journal of Orthopsychiatry*, *74*, 43–55.

Kromrey, J. D., & Foster-Johnson, L. (1996). Determining the efficacy of intervention: The use of effect sizes for data analysis in single-subject research. *Journal of Experimental Education*, *65*, 73–93.

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, 112–1134.

Leape, L. L. (1994). Error in medicine. *Journal of the American Medical Association*, *272*, 1851–1857.

Lichtenberger, E. O. (2006). Computer utilization and clinical judgment in psychological assessment reports. *Journal of Clinical Psychology*, *62*, 19–32.

Lilienfeld, S. O., Lynn, S. J., & Lohr, J. M. (2003). *Science and pseudoscience in clinical psychology*. New York: Guilford.

Lilienfeld, S. O., & O'Donohue, W. T. (2006). What are the great ideas of clinical science and why do we need them? *The General Psychologist*, *41*(2), 15–18.

Lonigan, C., Elbert, J. C., & Johnson, S. B. (1998). Empirically supported psychosocial interventions for children: An overview. *Journal of Clinical Child Psychology*, 27, 138–145.

Lopez, S. R. (1989). Patient variable biases in clinical judgment: Conceptual overview and methodological considerations. *Psychological Bulletin*, *106*, 184–203.

Macmann, G. M., & Barnett, D. W. (1999). Diagnostic decision making in school psychology: Understanding and coping with uncertainty. In C. R. Reynolds & T. B. Gutkin (Eds.), *Handbook of school psychology* (3rd ed., pp. 519–548). New York: Wiley.

Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology*, *34*, 362–379.

McCoy, S. A. (1976). Clinical judgments of normal childhood behavior. *Journal of Consulting and Clinical Psychology*, *44*, 710–714.

McDermott, P. A. (1980). Congruence and typology of diagnoses in school psychology: An empirical study. *Psychology in the Schools*, 17, 12–24.

McDermott, P. A. (1981). Sources of error in psychoeducational diagnosis of children. *Journal of School Psychology*, *19*, 31–44.

McDermott, P. A., Marston, N. C., & Stott, D. H. (1993). *Adjustment scales for children and adolescents*. Philadelphia, PA: Edumetric and Clinical Science.

McFall, R. M. (1991). Manifesto for a science of clinical psychology. *The Clinical Psychologist*, *44* (6), 75–88.

McFall, R. M. (2000). Elaborate reflections on a simple manifesto. *Applied & Preventive Psychology*, *9*, 5–21.

McFall, R. M. (2005). Theory and utility—Key themes in evidence-based assessment: Comment on the special section. *Psychological Assessment*, 17, 312–323.

McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology*, *50*, 215–241.

Meehl, P. E. (1960). The cognitive activity of the clinician. *American Psychologist*, *15*, 19–27.

Meehl, P. E. (1973). Why I do not attend case conferences. In P. E. Meehl, *Psychodiagnosis: Selected papers* (pp. 225–302). Minneapolis: University of Minnesota Press.

Meehl, P. E. (1997). Credentialed persons, credentialed knowledge. *Clinical Psychology: Science and Practice*, *4*, 91–98.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, *52*, 194–216.

Metz, C. E., & Pan, X. (1999). "Proper" binormal ROC curves: Theory and maximum-likelihood estimation. *Journal of Mathematical Psychology*, *43*, 1–33.

Moon, G. W., Fantuzzo, J. W., & Gorsuch, R. L. (1986). Teaching WAIS-R administration skills: Comparison of the mastery model to other existing clinical training modalities. *Professional Psychology: Research and Practice*, *17*, 31–35.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.

Nickerson, R. S. (2004). *Cognition and chance: The psychology of probabilistic reasoning*. Mahwah, NJ: Erlbaum.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and short-comings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, *39*, 84–97.

Palmiter, D. J. (2004). A survey of the assessment practices of child and adolescent clinicians. *American Journal of Orthopsychiatry*, *74*, 122–128.

Pappadopulos, E., Jensen, P. S., Schur, S. B., MacIntyre, J. C., Ketner, S. Van Orden, K., et al. (2002). 'Real world' atypical antipsychotic prescribing practices in public child and adolescent inpatient settings. *Schizophrenia Bulletin*, *28*, 111–121.

Paulos, J. A. (1998). *Once upon a number: The hidden mathematical logic of stories*. New York: Basic Books.

Persons, J. B., & Bertagnolli, A. (1999). Interrater reliability of cognitive-behavioral case formulations of depression: A replication. *Cognitive Therapy and Research*, *23*, 271–283.

Piattelli-Palmarini, M. (1994). *Inevitable illusions: How mistakes of reason rule our minds*. New York: Wiley.

Platt, J. R. (1964). Strong inference. *Science*, *146*, 347–352.

Plous, S. (1993). *The psychology of judgment and decision making*. New York: McGraw-Hill.

Poortinga, Y. H., & Soudijn, K. A. (2003). Ethical principles of the psychology profession and professional competencies. In W. O'Donohue & K. Ferguson (Eds.), *Handbook of professional ethics for psychologists: Issues, questions, and controversies* (pp. 67–80). Thousand Oaks, CA: Sage.

Popper, K. (1992). *In search of a better world: Lectures and essays from thirty years*. New York: Routledge.

Reason, J. (1990). *Human error*. New York: Cambridge University Press.

Robinson, J. O. (1998). *The psychology of visual illusion*. Mineola, NY: Dover Publications.

Rosenhan, D. L., & Messick, S. (1966). Affect and expectation. *Journal of Personality and Social Psychology*, *3*, 38–44.

Ruscio, J. (2003). Holistic judgment in clinical practice. *Scientific Review of Mental Health Practice*, *2*, 49–60.

Rutter, M., & Yule, W. (2002). *Child and adolescent psychiatry* (4th ed.). Malden, MA: Blackwell Science.

Sandoval, J. (1998). Critical thinking in test interpretation. In J. Sandoval, C. L. Frisby, K. F. Geisinger, J. D. Scheuneman, & J. R. Grenier (Eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 31–49). Washington, DC: American Psychological Association.

Schustack, M. W., & Sternber, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, *110*, 101–120.

Shear, M. K., Greeno, C., Kang, J., Ludewig, D., Frank, E., Swartz, M. D., & Hanekamp, M. (2000). Diagnosis of nonpsychotic patients in community clinics. *American Journal of Psychiatry*, *157*, 581–587.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118.

Skrabanek, P., & McCormick, J. (1990). *Follies and fallacies in medicine*. Buffalo, NY: Prometheus Books.

Slate, J. R., Jones, C. H., Coulter, C., & Covert, T. L. (1992). Practitioners' administration and scoring of the WISC-R: Evidence that we do err. *Journal of School Psychology*, *30*, 77–82.

Smith, D., & Dumont, F. (1997). Eliminating overconfidence in psychodiagnosis: Strategies for training and practice. *Clinical Psychology: Science and Practice*, *4*, 335–345.

Smith, J. D., & Dumont, F. (2002). Confidence in psychodiagnosis: What makes us so sure? *Clinical Psychology and Psychotherapy*, *9*, 292–298.

Strauss, G., Chassin, M., & Lock, J. (1995). Can experts agree when to hospitalize adolescents? *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 418–424.

Stricker, G. (2006). The local clinical scientist, evidence-based practice, and personality assessment. *Journal of Personality Assessment*, *86*, 4–9.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*, 1285–1293.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*, 1–26.

Tenopir, C., & King, D. W. (1997). Trends in scientific scholarly publishing in the United

States. *Journal of Scholarly Publishing*, *28*, 135–170.
Tracey, T. J., & Rounds, J. (1999). Inference and attribution errors in test interpretation. In J. W. Lichtenberg & R. K. Goodyear (Eds.), *Scientist-practitioner perspectives on test interpretation* (pp. 113–131). Boston: Allyn and Bacon.
Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*, 105–110.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.
Tversky, A., & Kahneman, D. (1993). Belief in the law of small numbers. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 341–349). Hillsdale, NJ: Erlbaum.
United States Department of Education. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide*. Retrieved March 1, 2006, from www.ed.gov/rschstat/research/pubs/rigorousevid/index.html.
Watkins, M. W. (2003). IQ subtest analysis: Clinical acumen or clinical illusion? *Scientific Review of Mental Health Practice*, *2*, 118–141.
Wechsler, D. (1991). *Wechsler intelligence scale for children—Third edition*. San Antonio, TX: The Psychological Corporation.
Wedding, D., & Faust, D. (1989). Clinical judgment and decision making in neuropsychology. *Archives of Clinical Neuropsychology*, *4*, 233–265.
Weinstein, C. S. (1988). Preservice teachers' expectations about the first year of teaching. *Teaching and Teacher Education*, *4*, 31–40.
Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*, 806–820.
Weinstein, N. D., Slovic, P., & Gibson, G. (2004). Accuracy and optimism in smokers' beliefs about quitting. *Nicotine and Tobacco Research*, *6*, 375–380.
Weinstein, S., Obuchowski, N. A., & Lieber, M. L. (2005). Clinical evaluation of diagnostic tests. *American Journal of Roentgenology*, *184*, 14–19.
Wiggins, J. S. (1988). *Personality and prediction: Principles of personality assessment*. Malabar, FL: Krieger Publishing Company.
Wilson, E. O. (1995). Science and ideology. *Academic Questions*, *8*(3), 73–81.
Ysseldyke, J. E., Algozzine, B., Regan, R., & McGue, M. (1981). The influence of test scores and naturally occurring pupil characteristics on psychoeducational decision making with children. *Journal of School Psychology*, *19*, 167–177.
Zarin, D. A., & Earls, F. (1993). Diagnostic decision making in psychiatry. *American Journal of Psychiatry*, *150*, 197–206.