CHAPTER 21

CLINICAL JUDGMENT AND DECISION-MAKING

Peter Salovey
Dennis C. Turk

Clinicians make judgments frequently about the possible causes and consequences of their clients' behaviors. They must decide on the best possible treatment for clients and try to predict how clients are likely to behave in the future. When prediction is based on intuitive, subjective, and idiosyncratic variables, it is often referred to as clinical judgment. Clinical judgment is often contrasted in reliability and validity with statistical or actuarial prediction, which is based on the quantitative combination of objectively gathered data.

The relative accuracy of statistical versus clinical predictions was a source of considerable debate after the publication of several highly influential articles by Sarbin (1941, 1943, 1944) and, especially, upon publication of Meehl's (1954) Clinical Versus Statistical Prediction, which summarized a number of studies pointing to the advantages of statistically based approaches to diagnosis and assessment. Nearly 35 years after publication of his book, Meehl (1986) claimed that no more than 5% of it needs to be retracted or in any way qualified. Even Holt, Meehl's most persistent critic (e.g., Holt, 1958, 1970, 1978), recently acknowledged that "maybe there are still lots of clinicians who believe that they can predict anything better than a suitably programmed computer; if so, I agree that it is not only foolish but at times unethical of them to do so" (Holt, 1986, p. 378). Given this seeming rapprochement between the clinical and statistical camps, it is probably most useful in a handbook such as this one to describe the cognitive processes that impinge on the validity of clinical judgment rather than to continue to polarize the issue as a debate about the advantages of objective, statistical prediction over intuitive, clinical judgment.

In this chapter, we will discuss research that is
relevant to the cognitive processes underlying the judgments of mental health professionals, including clinical psychologists. Our goal is not to demonstrate that clinicians are somehow deluded into believing that they possess especially acute judgmental skills and that they should perhaps be punished, or, better yet, hospitalized for these delusions. Rather, we hope to emphasize the basic notion that clinicians are decision makers just like any other decision makers, and their judgmental powers are probably no better and no worse than anyone else’s. Moreover, the manner by which a clinician makes a judgment or decision is identical to such processes in the lay person. This statement is not intended to be pejorative, but one of existing facts as we know them. Meehl (1954) underscored this point right from the outset of the clinical judgment debate when he noted that “psychologists should be sophisticated about errors of observing, recording, retaining, and recalling to which the human brain is subject. We, of all people, ought to be highly suspicious of ourselves. We have no right to assume that entering the clinic has resulted in some miraculous mutation and made us singularly free from ordinary errors” (pp. 27–28).

Meehl’s (1954) monograph summarized the literature on clinical judgment available at that time: 22 studies that could be reviewed in a brief chapter. Since the publication of Meehl’s volume, the literature has grown from 22 to perhaps more than 1000 articles directly relevant to clinical judgment and decision-making. There has been considerable speculation about the origins of the present interest in this area (Sarbin, 1986; Scriven, 1979). Kleinmuntz (1984) outlined five developments that have contributed to the recent scientific interest in clinical judgment. These include (a) lingering aspects of the original clinical versus statistical prediction controversy, (b) the growing interest in artificial intelligence and decision-making technologies among cognitive scientists, (c) the vast growth of behavioral decision theories and the decision sciences in the past 20 years, (d) a greater attention to formal approaches to medical reasoning among physicians, and (e) the emergence of the desktop computer as an information-processing machine.

In the remainder of this chapter, we will discuss some of the cognitive processes that affect the judgmental tasks of clinicians (see also Turk & Salovey, 1985; 1986). This description will be organized roughly in terms of how these tasks are sequenced in a typical clinical setting: (a) forming a first impression, (b) using a classification system, (c) making a diagnosis, (d) assessing covariation, and (e) making decisions and predicting the future. We will conclude with a discussion of how both clinical judgment and research on clinical judgment might be improved.

COGNITIVE PROCESSES IN CLINICAL JUDGMENT

Forming a First Impression

Prior Expectancies and Implicit Theories

Prior experiences, training, theoretical orientation, and influential authority figures often direct the expectations with which clinicians initiate an encounter with a potential client. When well organized, these sets of expectations can be referred to as schemas. Schemas play an important role in helping us filter and organize incoming information and attend primarily to important and relevant features of it. Without schemas, clinicians (or anyone else) would have difficulty making sense of the barrage of incoming information available in the clinical setting. Moreover, without schemas they would be unable to make decisions quickly.

These sets of expectations, however, can bias the processing of new information by focusing clinicians’ attention on information consistent with the schema and causing them to ignore or discount information irrelevant to these expectations. As a result, there is a tendency to process information confirming an initial expectation or impression (Bieri, Atkins, Briar, Leaman, Miller, & Tripodi, 1966; Hastie, 1981; Taylor & Crocker, 1981).

The potential biases fostered by the existence of prior expectations have been demonstrated in dramatic ways. Temerlin (1968) developed a tape that portrayed a man as happy and effective in his work, involved in warm and satisfying relationships, self-confident, and secure with little arrogance, competitiveness, or grandiosity. The man identified with his father, was happily married, and reported having pleasurable sexual experiences. He reported a happy childhood, had a good sense of humor, few role anxieties, and reasonable worries. In short, he displayed no ostensible symptoms of pathology and seemed to be a “normal, healthy man.”
Psychologists, psychiatrists, and graduate students were asked to observe the tape and determine whether this man was psychotic, neurotic, or healthy. In one condition of the experiment, a prestigious clinician introduced the tape by noting that the man portrayed was a "rare case of a mentally healthy individual." In another condition, he suggested that the man "looks neurotic but is actually quite psychotic." A control group received no suggestion at all. In the first condition, all of the clinicians rating the tape agreed that the man was in fact a healthy individual. However, in the condition during which they were told to expect "latent" psychosis, 44% of the clinicians evaluated the man as psychotic, 50% as neurotic, and only 6% as healthy. When no suggestion at all was given, 57% of the clinicians thought the man was healthy, none of the clinicians rated the man as psychotic, although 43% thought he was neurotic. The prior expectation of psychosis (or, at least, "looking neurotic") in the psychosis suggestion condition may have caused the clinicians to pay particular attention to and remember evidence consistent with pathology. The bias was most strongly held by psychiatrists and least strongly held by graduate students of clinical psychology. Professional clinical psychologists were moderately influenced by the prestigious psychotherapist.

In another study, when a taped actor was labeled a "job applicant," he was seen as much less psychologically disturbed than when the identical behavior was portrayed but the actor labeled a "patient" (Langer & Abelson, 1974). The impact of this prior expectation was stronger for psychodynamically oriented clinicians than for behaviorists. Perhaps the behaviorists attended only to the patient's objectively observable behaviors (but see Snyder, 1977, for an alternative explanation.) Other experiments like these have been conducted and the effects observed by Tenerlin, and Langer and Abelson have largely been replicated (e.g., Sushinsky & Wener, 1975). However, this line of research has been criticized on theoretical grounds (e.g., Davis, 1979, who argues that the clinicians in these experiments are appropriately utilizing base-rate information).

Even when prior expectancies are not explicitly provided, a set of expectations may develop rapidly after initial patient contact, often based on very little information. For example, after determining that someone is intelligent, we may expect him or her to have a good sense of humor even though we have never seen any evidence of wittiness. These implicit personality theories (Asch, 1946; Bruner & Tagiuri, 1954) are then used as the basis for subsequent judgments about the individual even if the clinician no longer can recall the origins of the implicit theory itself (Carlston, 1977; Srull & Wyer, 1979). Unfortunately, individuals adhere tenaciously to initial diagnostic conceptualizations, largely ignoring new information (Rubin & Shontz, 1960).

The clinicians in Rosenhan's (1973) classic study, in which normal individuals falsified a single psychotic symptom (hearing voices), readily diagnosed these pseudopatients as pathological (usually schizophrenic) and admitted them as inpatients for 7 to 52 days in the absence of any subsequent evidence of psychosis (see Spitzer, 1975, 1976; Weiner, 1975, for critiques, however). The initial impression of pathology made it difficult for these clinicians to recognize "normal" behavior in the pseudopatients. Meehl (1960) observed that the basic impression of a patient formed during the first few sessions is retained intact after 24 sessions, and the initial diagnostic impression of a patient may be formed in the first minute of therapy (Gauron & Dickinson, 1969; see also Sandifer, Hordern, & Green, 1970). This diagnostic tenacity may be in part a result of tendencies to seek or elicit information confirming initial hypotheses, a judgmental phenomenon that will be dealt with in the next section of this chapter.

**Confirmatory Strategies**

Prior expectations, schemas, and implicit theories all affect initial impressions, as we have described. Additionally, they profoundly influence subsequent data gathering. When these expectations are organized into hypotheses, there is a pervasive tendency to seek information that confirms them and to ignore disconfirming information. This data-gathering style has been labeled "behavioral confirmation" or "seek and ye shall find" (Snyder, 1981). Snyder defines behavioral confirmation as the process by which an individual's preconceived beliefs and prior expectations guide interaction in such a way that these initial beliefs, even when false, come to be confirmed by the other person's behavior (Snyder & Thomsen, 1988).

Individuals tend to use their interactions with others as opportunities to test the accuracy of hypotheses they hold about them. But we are not objective observers of them. Rather, we often act as if the hypotheses are true, eliciting hypothesis-
confirming behavior from others. Some years ago, Merton (1948) referred to such an interaction sequence as a self-fulfilling prophecy. For example, when college students were asked to find out whether another person was an extrovert or an introvert, they selected questions based on whether the initially provided hypothesis was one of introversion or extroversion (Snyder & Swann, 1978b). If they were trying to confirm whether the person was an extrovert, they would ask questions that solicited evidence of extroverted behavior (e.g., “What would you do if you wanted to liven things up at a party?”). If they were trying to confirm the introvert hypothesis, they might select a question such as, “What factors make it hard for you to really open up to people?” Not only do these questions elicit a biased set of information, but they may cause the other person to behave in a way that confirms the hypothesis (e.g., one is more likely to behave in an outgoing way while discussing how to liven up a party as compared with discussing what makes it hard to open up to others).

The pervasiveness of the hypothesis-confirmation strategy is quite breathtaking. Individuals engage in this information-seeking style no matter what the origins of the initial hypothesis (e.g., self-generated, provided by others), whether the hypothesis is likely to be accurate or inaccurate, whether clear incentives for accurate hypothesis testing are offered, or whether the hypothesis contains disconfirming attributes (Snyder, 1981). Surprisingly, subjects in Snyder’s experiments treated all hypotheses with equal weight when selecting questions to ask, and these hypothesis-confirming strategies even generalized to drawing inferences about oneself.

The seeking of hypothesis-confirming information may create a tendency among clinicians to overpathologize the behavior of others. Because in the clinical context clinicians expect to observe pathology, there may be a pervasive tendency to interpret ambiguous client behaviors as evidence for pathology (Sarbin, Taft, & Baily, 1960). Another danger of hypothesis-confirming strategies in the clinical setting is that clinicians at times are motivated to elicit experiences from their clients that confirm their theoretically based hypotheses. After a time in therapy, clients’ behaviors may come to match the theoretical frameworks of their therapists (Frank, 1974; Scheff, 1966). In fact, a therapeutic cure might be defined as the conversion of the client’s value and belief systems to that of the therapist (Bandura, Lipsher, & Miller, 1960; Rosenthal, 1955; Welkowitz, Cohen, & Ortmeyer, 1967). After a time in psychotherapy, clients will even report dream material that is consistent with the therapist’s orientation (Whitman, Kramer, & Baldridge, 1963).

Salience, Vividness, and Availability

Another variable influencing the forming of an initial impression is the notion that some people and some features of other people are more noticeable than others. Salience refers to the properties of a stimulus that make it likely to attract attention relative to its context. Vividness is often used to denote the inherent properties of a stimulus that attract attention independent of its context (Fiske & Taylor, 1984). People become salient when they are novel in a social context (e.g., the only blond in a room full of brunets or the only schizophrenic in a room of “normals”), by behaving in unexpected ways, or by being relevant to the perceiver’s goals. On the other hand, a description of a specific traffic accident is more vivid than statistics about traffic accidents, no matter the context in which they are presented.

It might be expected that particularly salient or vivid aspects of an individual would exert greater impact on judgment. That is, the mundane is forgotten while the unusual remembered. Salient individuals are seen as more in personal control of their behavior rather than as buffeted by external forces (Fiske, Kenny, & Taylor, 1982). When a person acts especially crazy among a crowd of people acting normally, we assume that the crazy person is responsible for his or her odd behavior, discounting the role of environmental factors. Moreover, unpleasant salient others are likely to be judged as especially unpleasant (Fiske & Taylor, 1984). Surprisingly, though, these attributional and evaluative consequences of salience may not be mediated by memorability (McArthur, 1981; Taylor & Fiske, 1978; Taylor & Thompson, 1982). That is, salience may influence attributions and evaluations even when there is no evidence of increased recall of salient others.

On those occasions, however, when especially salient or vivid material is encoded better into memory, it can exert a pronounced and systematic effect on judgment. Information that is especially easy to bring to mind is said to be more available. When individuals estimate the probability of an event or outcome by the ease with which it is brought to mind, they are using the availability
heuristic (Tversky & Kahneman, 1973). When forming an impression of a new client, it is easy to overestimate the frequency of the client's unusual behaviors because such behaviors are more available. For example, a clinician might decide that a client is suicidal because the client once mentioned thinking about suicide. The memorability of the client's revelation leads to an overestimation of its likelihood.

Another example concerns judgments about the dangerousness of released mental patients. Often, the dangerousness of discharged mental patients is overestimated because the occasional violent patient is easily recalled, with help from media coverage, but the more typical expatiant who lives out a quiet and uneventful life is not remembered. Hence, a clinician easily recalling a dramatically violent former patient may form an impression of a new client as being more dangerous than he or she actually is. Similarly, we may overestimate the proportion of the homeless who are mentally ill because “normal” acting homeless individuals are not especially memorable.

**Emotional Influences on Person Perception**

A final variable that influences initial impressions is affect. The moods and emotions of the perceiver influence judgments made about the characteristics of other people. For example, whether one's mood is created by a recent failure (Lerner, 1965), receipt of a desirable prize (Lott & Lott, 1968), viewing of an evocative film (Gouaux, 1971), or ambient temperature (Griffitt, 1970), we find others to be more attractive and desirable when we are in pleasant rather than dysphoric moods ourselves.

When incoming information is ambiguous, moods may be especially likely to bias initial impressions. For example, happy individuals are more likely to rate the illustration on a Thematic Apperception Test (TAT) card as containing pleasant content than are angry individuals (Bower, 1981), and they are more likely to interpret ambiguous written material in a similar mood-congruent manner (Clore, Schwarz, & Kirsch, 1983).

In general, judgments about others follow a mood-congruent pattern (Forgas & Bower, 1988). When people are in good moods, their judgments are positive, and when in bad moods, negative (Mayer & Bremer, 1985; Mayer & Volanth, 1985). One reason why moods might affect judgment is that they make mood-congruent events stored in memory more easily retrievable, and the greater availability of such memories leads to predictions of greater likelihood for similar future events (Bower, 1981; Isen, 1970, 1984). An alternative explanation (Mayer, 1986) is that such judgment effects are rooted in recategorizations in memory rather than memory enhancements or failures. According to this view, individuals in a good mood may classify a mildly negative event as a neutral or positive event (Mayer & Salovey, 1988).

**Using a Classification System**

Clinical judgment in psychology is often based on a taxonomy. In order to communicate, conduct research, and prescribe treatment, it is essential that some consensually validated criteria are used to describe groups of individuals who are similar on some relevant attributes. For clinical psychologists, the most recent edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III-R; American Psychiatric Association, 1987) serves as the guiding taxonomic framework for the ultimate classification of individuals into diagnostic categories. The construction of a classification system and its eventual use are psychological processes, no matter how objective they may seem, and it is to these issues that we will turn briefly.

**Development and Use of a Taxonomy**

A taxonomy of taxonomies would divide them into either natural or artificial and hierarchic or nonhierarchic (Rasch, 1987). When the criteria used to classify individuals (as in the case of psychologists) are based on characteristics fundamental to those individuals, the taxonomy is termed natural. Artificial taxonomies classify individuals on the basis of an accidental or arbitrary attribute. A taxonomy based on height or eye color clearly would be a natural one, whereas one based on shirt color would be artificial. The DSM-III-R was intended to be a natural taxonomy, purportedly describing differences inherent within individuals. Critics, however, who argue that DSM-III-R diagnostic categories lie primarily in the minds of the diagnostician rather than in the observable attributes of clients and patients might claim that the DSM-III-R is actually an artificial taxonomy (cf. Rosenhan, 1975). That is, they would claim that the defining criteria "reside" in the observer and his or her context, not in the observed.

Hierarchic taxonomies attempt to group indi-
viduals into successively more inclusive classes based on fewer attributes to define a particular category. Nonhierarchic taxonomies use a single attribute or set of attributes to define a category (Lowry, 1981). The DSM-III-R is largely non-hierarchic, but it does contain some hierarchical elements. For example, in the DSM-III-R, the category of schizophrenic disorders is subdivided further into disorganized, catatonic, paranoid, and residual types.

Taxonomies may be developed inductively or deductively (Mezzich, 1980; Walker & Avant, 1983). Some classification schemes in psychology are developed inductively or empirically. That is, individuals are observed until some attribute on which they can be differentiated is discovered. Individual differences in the members of these categories are then ascertained such that the categories can be divided into sets of subcategories. After individuals are categorized, a taxonomy can be developed that specifies the rules and procedures for further classification of individuals. These rules were thus derived inductively by observing differences among individuals.

A deductively derived taxonomy begins with a theoretically based statement about how individuals should differ and subsequently be categorized. Its utility is then tested by actually employing it to classify individuals. Does assignment of an individual to a class facilitate treatment decisions or predictions of future behavior? It is probably accurate to say that the major categories of the DSM-III-R (e.g., schizophrenic disorders, organic mental disorders, affective disorders, anxiety disorders) were derived deductively, and that the specific subtypes within these general categories were constructed inductively.

Several structural issues, reviewed by Mezzich, Goodpastor, and Mezzich (1987), constrain the ways in which clinicians make clinical judgments in response to a diagnostic taxonomy. The first of these issues is that of polydiagnosis. All psychiatric disorder criteria contain exclusion criteria when making a diagnosis. This polydiagnostic approach would allow individuals to carry diagnoses based on different taxonomies. Although potentially confusing, the advantage would be a clear statement of a patient's diagnosis according to each of a set of taxonomies (e.g., DSM-III-R, research diagnostic criteria, and Bleuler's [1908] standard). The argument for polydiagnosis also includes the notion that systematic discrepancies among diagnostic systems would be more easily identified and ameliorated. In clinical research, polydiagnosis would increase the comparability of results across research teams using different diagnostic criteria.

The second of these structural issues is the hierarchical arrangement of clinical disorders such that the diagnosis of one precludes the diagnosis of another. In the most parsimonious hierarchy, no one individual could be diagnosed with more than one disorder (Jaspers, 1963). In fact, when most taxonomies are used, such as the DSM-III-R, the presence of one disorder increases the probability that a second disorder might be diagnosed. For DSM-III-R, the receipt of any diagnosis increases the probability of receiving almost any other diagnosis. Mezzich et al. (1987) suggest that explicit exclusion criteria might be developed in order to minimize inaccuracies resulting from such category overlap. At present, only some disorders listed in the DSM-III-R contain exclusion criteria.

A third structural issue is determining how multiple diagnoses should be integrated when the criteria for several different disorders are met. For example, a patient might be diagnosed as having a simple phobia, being an abuser of alcohol, and behaving like a depressive. Such multiple diagnoses can be handled in a variety of ways. They can be combined into syndromes when their co-occurrence is systematic. Alternatively, the person can be treated as if the diagnoses were conceptually (and etiologically) distinct.

Finally, Mezzich et al. (1987) urged us to consider multiaxial diagnosis in which several different aspects of a psychological disorder are considered systematically. A multiaxial system asks the clinician to make a diagnostic judgment across a set of aspects (termed axes). The DSM-III-R introduced multiaxial diagnosis to most clinicians. Judgments about patient diagnoses are now made on five (not necessarily orthogonal) dimensions: primary psychiatric syndromes, personality and developmental disorders, physical disorders, psy-
chosocial stressors and their severity, and highest level of adaptive functioning during the past year. Use of the third (physical illness), fourth (stress), and fifth (level of functioning) axes have been promoted, especially, by the health psychology movement in clinical psychology. In practice, however, often only the first and sometimes the second axis are used. Limits on the use of the full multiaxial system are often imposed by (a) insurance companies who demand that patients be labeled on the first axis primarily, (b) clinicians who are trained to accentuate this pathology dimension, (c) diagnosticians who may have difficulty simultaneously conceptualizing clients on a large number of dimensions, and (d) the time and financial pressures on the practicing clinician that reward quick and parsimonious decisions.

Diagnostic Categories and the Processing of Prototypes

Cantor and her colleagues (e.g., Cantor, Smith, French, & Mezzich, 1980; Genero & Cantor, 1987) have noted that diagnostic categories based on taxonomies such as the DSM-III-R have some special properties. First, it is difficult to specify the defining features that all category members possess. Further, these categories (i.e., diagnoses) are vague or “fuzzy” (cf. McCloskey & Glucksberg, 1978); that is, they sometimes lack consensually defined and specific necessary and sufficient criterial properties. As a result, diagnostic categories may be better defined by specific prototypes rather than criterial features (Horowitz, Post, French, Wallis, & Siegelman, 1981; Horowitz, Wright, Lowenstein, & Parad, 1981). In other words, diagnosticians may think about a category such as “paranoid schizophrenic” by retaining in memory a prototypic paranoid schizophrenic rather than a list of features common to all paranoid schizophrenics.

The implications of the idea that diagnostic categories are fuzzy sets best characterized by prototypes who possess a set of correlated features are that there is considerable heterogeneity among similarly diagnosed patients, that similarly diagnosed patients can vary in terms of how typical they are for (i.e., how well they fit) a particular diagnosis, and that diagnosis is the process by which a clinician matches the features presented by a patient with those of category prototypes (Genero & Cantor, 1987). The diagnostic process, then, is characterized by a similarity-matching procedure whereby a particular client is compared with a prototype for each potentially relevant diagnostic category. This is a very different process than simply affirming the existence of a small group of requisite features. As a result, some clients will be more easily classified and subsequently remembered than others—those clients who most closely match the category prototype. Imagine that each category is a circle. The exact center of the circle depicts the prototype for people who fit in the circle (i.e., the category). However, an individual can be closer to the periphery than the center of the circle yet still have enough of the defining characteristics to be classified within the category depicted by the circle. The difficulties of diagnosis by prototype-matching will be discussed momentarily.

Making a Diagnosis

After forming an initial impression of a client, clinicians often must get down to the business of using a taxonomy to make a formal diagnosis. Although many of the same cognitive processes operate for this task as were described in initial impression formation, the process of diagnosis also lends itself to other judgmental difficulties.

Prototype-Matching

As described earlier, clinical diagnosis proceeds through prototype-matching, whereby a patient’s characteristics are compared with those of a category prototype (Genero & Cantor, 1987). This diagnostic process may pose some judgmental difficulties. Because diagnostic categories are fuzzy, the individuals receiving a particular diagnosis are a heterogeneous group. Hence, it can be difficult to recognize even typical category members. For example, among a group of individuals diagnosed as depressed, there may be individuals with different subsets of depressive characteristics. One person might be agitated, another lethargic, and so forth. Moreover, even clients who match a category prototype quite closely may also bare some resemblance to another category prototype. For example, a paranoid schizophrenic may also seem somewhat like a person with bipolar affective disorder (i.e., manic-depressive psychosis) in a manic phase.

Representative Thinking

The diagnostic process is essentially a judgment that an individual is a member of a particular category or that a given outcome can be explained
by a particular set of antecedents. The decision-making strategy frequently employed for judgments of this kind is the representativeness heuristic (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1974). Representativeness refers to judgments based on the degree to which a given stimulus or evidentiary base matches the essential features of some category, schema, or prior expectation. To make an accurate decision, clinicians must consider the probability of encountering a category member as well as a nonmember by chance, given that both exhibit some diagnostic sign. However, much of this important base-rate information is ignored, and decision-makers tend to attend selectively instead to information that fits preexisting sets of expectations, as we discussed previously.

It is the representativeness heuristic that makes us think that a bright college student who is a poor speller might be dyslexic, even though there are innumerable individuals who are not dyslexic but are still poor spellers (Dawes, 1986). We are ignoring the high base-rate of poor-spelling non-dyslexics in the population and instead focusing our attention on a presumed link between spelling and dyslexia. As a general rule, we tend to ignore the base-rate of a particular disorder independent of some characteristic of the disorder, and vice versa (Dawes, 1963; Meehl & Rosen, 1955). We may believe that someone giving a color response on the Rorschach is emotionally labile, ignoring the number of people who are emotional who do not give color responses as well as the number of people who do not give color responses but are emotionally labile.

Use of the representativeness heuristic gives rise to a classic judgmental error called illusory correlation, the perception that events are associated even when the relationship between them is incidental (Chapman, 1967; Tversky & Kahneman, 1980). In a classic set of studies, for example, clinicians readily assumed that certain signs on projective drawing tests (e.g., the Draw a Person Test) are associated with clinically relevant patient characteristics, even in the absence of any demonstrated association. For example, the vast majority of clinicians assumed that drawings containing unusual eyes indicate suspiciousness and that muscular physiques indicate a concern with masculinity (Chapman & Chapman, 1967). Interestingly, untrained subjects espoused these same illusory associations. Again, the actual base-rates of the symptom (suspiciousness) in the presence or absence of the sign (unusually drawn eyes) are ignored, as is the base-rate of perfectly normal people drawing unusual eyes when presented with this task.

Similar results have been obtained using the Rorschach inkbloths. Diagnosticians consistently overestimate the association between popular responses (e.g., seeing an anus) and stereotypically associated conditions (e.g., homosexuality) (Chapman & Chapman, 1969). Illusory correlations have been demonstrated using behavioral rating systems as well (Berman & Kenny, 1976). The bias is especially tenacious. Even after being told that correlations are illusory or being taught alternative diagnostic strategies, diagnosticians still adhere to the illusory associations. In fact, they tend to become even more confident of their veracity (Einhorn & Hogarth, 1978; Kurtz & Garfield, 1978; Mowrey, Doherty, & Keeley, 1979; Waller & Keeley, 1978). We consider such illusory correlations as the result of representative thinking. However, the problem of illusory correlation extends beyond this domain and will be dealt with in its own right in the next section of this paper.

Assessing Covariation

Many clinical judgment tasks, at all phases of diagnosis and treatment, involve covariation assessment, the determination of a relationship between two events; in particular, that one event occurs more in the presence than in the absence of the other event (Kayne & Alloy, 1988). As Kayne and Alloy described, much of the therapeutic process is taken up with assessing covariation. For instance, some of the biases that we have discussed under other rubrics can also be considered as instances of covariation assessment gone awry. The illusory correlations reported by Chapman and Chapman (1967, 1969) and discussed by us as an example of representative thinking can be considered a failure in accurately assessing the covariation between psychodiagnostic test results and patients' symptoms. Similarly, the confirmatory biases we have discussed, such as the labeling bias demonstrated by Teverlin's (1968), Rosenhan's (1973), and Langer and Abelson's (1974) subjects, can all be considered distortions in the assessment of actual covariation between behaviors implied by such labels and the patient's behavior. Difficulties in assessing covariation also tend to bias clinicians to be overconfident about the likely success of treatments consistent with their schooling.
Errors in Detecting Covariation

The literature on the ability to detect covariation accurately has been well described by Kayne and Alloy (1988), so we will only summarize it briefly here. In short, if individuals are asked to judge the covariation of sets of continuous stimuli about which they have no prior expectations (e.g., lists of number pairs), they can do it quite accurately (Beach & Scopp, 1966; Erlick & Mills, 1967). However, when asked to make dichotomous judgments about stimuli for which individuals have strong prior expectancies, the evidence for accuracy is weaker (Alloy & Abramson, 1979; Dickinson, Shanks, & Evenden, 1983; Peterson, 1980; Ward & Jenkins, 1965). In these studies, typically, when subjects expected to observe covariation they in fact did, unless encouraged to consider noncontingency as an alternative hypothesis. Finally, when judging one's own behaviors, there is a marked tendency, so long as we are not depressed (see Abramson & Alloy, 1981; Alloy & Abramson, 1979, 1982; Alloy, Abramson, & Viscusi, 1981; Alloy, Crocker, & Tabachnik, 1980), to overestimate the degree of covariation between our actions and environmental outcomes (Langer, 1975; Langer & Roth, 1975; Wortman, 1975).

The Sources of These Errors

Why do individuals have difficulty accurately assessing covariation, especially when they have strong expectations about contingencies among events? Kayne and Alloy (1988) suggested that there are five points in the judgmental process from which biases in the assessment of covariation can emanate: (a) gathering confirming and disconfirming data, (b) sampling cases, (c) classifying instances, (d) recalling evidence and estimating frequencies, and (e) combining evidence and formulating a judgment. Suppose a clinician is trying to decide whether clients who are late for their first appointments tend to show little improvement in therapy. This clinician, then, is trying to assess the covariation between being late and improvement and expects to see a negative association. In order to assess whether there is a relationship between these two variables, the clinician would have to think about four different kinds of clients: those who were late and those who were not late crossed with those who improved and those who did not improve. As discussed earlier, however, individuals are more prone to attend to and remember cases that confirm the late/no-improvement, not-late/improvement hypotheses and to ignore the disconfirming cases that could be found in the other two cells (Snyder & Cantor, 1979; Snyder & Swann, 1978a, 1978b; Wason & Johnson-Laird, 1972). For the most part, people are more likely to attend to cases that confirm the primary hypothesis; in this case, the late/no-improvement combination (Crocker, 1982; Schustack & Sternberg, 1981).

The second point at which error is introduced into the assessment of covariance is when deciding on which cases to include as evidence. First, when the clinician is seeking to test the late/no-improvement hypothesis, he or she can easily forget that the available data base (i.e., that clinician's case load) is neither a random nor especially large sample of cases (Tversky & Kahneman, 1971). Moreover, if the clinician tries to enlarge the sample by asking colleagues, the question is likely to be formulated as, "Do you have any patients who came late to their first session and then never improved?" forgetting about the three other possible combinations of these variables, especially the two that disconfirm the expected association (Alloy, Crocker, & Tabachnik, 1980; John, Scott, & Bettman, 1986).

Individuals also have great difficulty correctly classifying cases as confirming or as disconfirming the initial covariance hypothesis. In the example that we have been using, what is considered a "late" client? One who arrived 5 minutes late? Ten minutes late? It is quite possible that a client who arrived 5 minutes late and subsequently did not improve might inadvertently be classified as "late" at the same time that a 10-minute-late client who was "cured" was not (Crocker, 1981). The subjective classification of clients as improved or not improved is probably even more difficult. The initial hypothesis about covariation is likely to influence the manner in which ambiguous cases are classified (Nisbett & Ross, 1980).

After classifying cases as confirming or disconfirming the covariation hypothesis, individuals must total them up. If the cases are aggregated over a long period of time, recall biases may make covariation estimation inaccurate. As mentioned earlier, cases consistent with initial hypotheses are more likely to be recalled than irrelevant or inconsistent cases (Arkes & Harkness, 1983; Trolier & Hamilton, 1986).

Once the cases in the four cells are recalled and combined, individuals must somehow arrive at an estimate of the actual level of covariation between
CLINICAL JUDGMENT

Making Decisions and Predicting the Future

Decision-Making Processes

The kind of decision-making faced by clinicians is generally referred to as decision-making under conditions of uncertainty. The study of such decision-making processes is called decision analysis (Elstein, 1988). Decision analysis is the examination of normative processes in decision-making (Keeney, 1982; Raiffa, 1968; Weinstein & Fineberg, 1980), especially logical, rational decisions that follow from the application of axiomatic principles such as the theory of expected utility (Schoemaker, 1980; Von Neumann & Morgenstein, 1947). However, psychological research on decision-making suggests that people often deviate from the decision-making outcomes predicted by decision theorists and expected utility theory (Fischhoff, 1980; Kahneman & Tversky, 1979; Tversky & Kahneman, 1981). Recently, Tversky and Kahneman (1986) argued that actual human decision-making processes will never consistently fit normative models. Rather, normative models, like expected utility, are ideals rather than actual representations of human decision-making processes (for an excellent review of decision analysis and decision theory, see Abelson & Levi, 1985).

Elstein (1988) nicely summarized the ways in which decision-makers deviate typically from normative or ideal models of decision-making. The first difficulty is in estimating probabilities for clinically relevant outcomes. Normatively, the determination of the probability of a particular clinical outcome (e.g., depression, schizophrenia) given a diagnostic sign is based on Bayes's Theorem, which instructs the decision-maker to attend to (a) the frequency that the diagnostic sign occurs in the population, (b) the frequency or base-rate of the clinical outcome in the population, (c) the probability that a person with the clinical outcome will display the diagnostic sign, and (d) the probability that normal individuals will display the diagnostic sign, as we have discussed earlier. However, individuals deviate from Bayes's Theorem considerably. Clinical decision-makers often attend only to a test's sensitivity (its accuracy in identifying individuals with a disorder) and ignore its specificity (its accuracy in identifying people who do not have the disorder) (Balla, Elstein, & Gates, 1983). They ignore the base-rates and attend to a few vivid cases (Nisbett, Borgida, Crandall, & Reed, 1976; Nisbett & Ross, 1980). Even if a diagnostic sign displays few false positives, when the base-rate of a disorder in a population is small, many people will be falsely labeled as pathological (Wiggins, 1973).

A second deviation of human decision-makers from normative decision theories reviewed by Elstein (1988) concerns the minimization of maximum possible losses (i.e., the minimization of regret, Feinstein, 1985). Clinicians will often make decisions that minimize possible aversive outcomes rather than ones that might maximize health. For example, clinicians might be inclined to overdiagnose pathology because that would minimize the risk of a grossly impaired person being labeled healthy and not receiving any treatment. Often such decisions are accompanied by an overconfidence in clinicians' beliefs about the salubriousness of their therapeutic actions (Cohen & Oyster-Nelson, 1981; Kayne & Alloy, 1988).

Finally, the estimates of the utility of various therapeutic options varies widely depending on how these options are framed (Tversky & Kahneman, 1981). For example, clinicians often come to different decisions if they focus on the probability of a positive outcome if an action is taken versus the probability of a negative outcome if an action is not taken. The framing postulate suggests that decision-makers organize information relevant to decisions in terms of potential gains or potential losses as compared with a present referent point (e.g., present level of health). Factually equivalent material can be presented to individuals such that they encode it as either a gain or a loss. For example, in a classic Tversky and Kahneman (1981) study, subjects were presented with a situation in which the outbreak of a disease is expected to kill 600 people. Subjects must decide whether to endorse a program guaranteeing that 200 people will be saved and 400 will die or one that says there is a .33 probability that all 600 will be saved and a .67 probability that all 600 will die. When the question emphasizes the fact that the first program will save 200 people, subjects tend to endorse this
“sure thing” option. But when the question emphasizes that this option will result in 400 deaths, subjects tend to endorse the risky, probabilistic alternative.

Affective Influences on Risk-Taking and Decisions

Earlier we reviewed some of the evidence for the idea that moods and emotions influence judgment in a direction congruent with their affective valence. In addition, affect seems to influence decision-making under conditions of risk or uncertainty (reviewed by Isen, 1987). For example, when the outcome of a decision is not especially important, happy moods tend to produce riskier decision-making. However, when there is a potential for a great loss, individuals tend to become more conservative when in a positive mood (Arkes, Herren, & Isen, 1988; Isen & Geva, 1987; Isen, Means, Patrick, & Nowicki, 1982; Isen & Patrick, 1983). When the risk of loss is great, happy subjects seem to behave so as to maintain their happiness. Happy individuals view the consequences of such losses as more dire or extreme than individuals in a more neutral mood state (Isen, Nygren, & Ashby, 1985). Moreover, moods have a rather systematic influence on perceptions of the probability of positive and negative outcomes. The probability of future negative events such as diseases, natural disasters, and other catastrophes looms much larger when we are sad than happy (Johnson & Tversky, 1983; Salovey & Birnbaum, 1989).

The decision-making strategies used by individuals vary depending on mood. Happy individuals are likely to try to reduce the complexity of a decision-making task in order to engage in quick and simple kinds of cognitive processing (Isen & Daubman, 1984). Moreover, when happy, individuals are more likely to use intuitive (and potentially error-prone) strategies as compared with more taxing, logical ones (Isen et al., 1982). As we have described elsewhere (Salovey & Turk, 1988), it seems that happy decision-makers may be more inclined to rely on their “gut” instincts.

The Hindsight Bias

After a decision is made, clinicians can be influenced by the tendency of their colleagues to claim that they could have predicted the outcome in advance. Labeled the hindsight bias by Fischhoff (1975), individuals believe that certain outcomes are easier to predict when thinking about them after the fact. Such “Monday morning quarter-

HUMAN INFORMATION-PROCESSING LIMITATIONS

The bulk of this chapter thus far has been devoted to the specific ways in which clinical judgment can be suboptimal. We have tried to emphasize that it is judgment per se that is fallible, not necessarily clinicians. In this section, we will discuss those aspects of the human cognitive system that limit judgmental capacities in all of us.

Faust (1986) has suggested that two cognitive limitations may be the major contributors to judgmental difficulties: the capacity to use additional information and the ability to perform multiple cue tasks. Individuals have great difficulty integrating new, additional information within an existing set of information. As additional information is provided, it is often ignored or discounted (Sines, 1959). For example, clinicians have great difficulty integrating new test results with an existing battery of scores (Golden, 1964) or with data gleaned from a clinical interview (Winch & More, 1956).

Second, clinicians often believe they are integrating various sources of data in a complex, interactional, configural analysis. However, in reality, studies using multiple cue tasks that provide information pertaining to two or more dimensions indicate that clinicians actually combine information in simple linear ways (Wiggins & Hoffman, 1968). This is probably why linear models provide such an accurate representation of human judgment processes (Goldberg, 1968). In short, individuals have great difficulty combining multiple cues even when the judgmental task requires the integration of only two or three bits of information. We seem to lack the cognitive capacity for such operations (for a more thorough review see
Slovic & Lichtenstein, 1971). These limitations on cognitive processing reflect what has been termed bounded rationality (Newell & Simon, 1972; Simon, 1979), the idea that humans possess a limited capacity for rational thought. Bounded rationality is primarily attributable to the limits of human memory so that complex problems must be somehow simplified in order to reduce their load on the cognitive system.

IMPROVING CLINICAL JUDGMENT

One of the reasons why the literature on bias in clinical judgment and inference processes has had little impact on clinical practice is that often such limitations are described in a derogatory tone and few hints about their reduction are provided. With great amusement, investigators in this field have documented our judgmental shortcomings and other "irrational" inferential processes. But it is much easier to be critical than constructive. Behavior is difficult to change, and sometimes the suggestions of nonpractitioners are overly cumbersome. For discussions of clinical judgment to have impact in clinical settings, however, they must be provided in the context of practical as well as helpful advice. In the remainder of this chapter, we will discuss some of the procedures that have been suggested for improving clinical judgment and some ideas for improving research in this area as well.

Direct Training in Reasoning Skills

It is often thought that if psychological and medical decision-makers could be instructed directly in judgmental skills and warned against inferential limitations and biases, they could learn to be more accurate and less overconfident judges. For example, Howe, Holmes, and Elstein (1984) offered a three-term course in decision analysis to medical students organized around the discussion of cases. Although the students often found presentations about the limitations of clinical judgment unsettling, they were able to learn to recognize and quantify sources of uncertainty and error.

The impact of such direct training programs on judgment has been assessed, but the results are mixed. Individuals do not seem to have difficulty learning abstract (not domain-specific) reasoning skills (Nisbett, Fong, Lehman, & Cheng, 1987). Moreover, statistics courses do increase the chances that individuals will recognize the operation of statistical principles in common events (Fong, Krantz, & Nisbett, 1986). However, the direct provision of corrective feedback to clinical decision-makers has not proved to be especially effective (Goldberg, 1968; Graham, 1971), nor have warnings provided in advance of decision-making (Fischhoff, 1977; Kurtz & Garfield, 1978; Wood, 1978).

General Correctives

Although merely pointing out the existence of judgmental shortcomings may not be an especially effective way in which to reduce their impact, a certain amount of self-reflection might be helpful. Faust (1986) described three general strategies that yield easily implemented corrective procedures. These include (a) testing the validity of diagnostic signs, (b) using disconfirmatory strategies, and (c) recognizing predictive uncertainty. To these we add, as others have advised (e.g., Dawes, 1986), suggestions regarding how to avoid being influenced by client acceptance of faculty judgments and how to reduce the burden on memory inherent in judgmental tasks.

Tests for Diagnostic Signs

Faust (1986) recommended that clinicians examine supposedly diagnostic tests and signs very carefully for evidence of their validity. In particular, three tests should be applied to a diagnostic sign before it is accepted as useful. The first test is to determine if there is a true association between the sign and the disorder. The disorder must appear more frequently in the presence of the sign than in its absence. As discussed earlier, the probability of finding a clinical case in each of four cells must be evaluated: (+)sign/(+)disorder, (+)sign/(-)disorder, (-)sign/(+)disorder, (-)sign/(-)disorder. Validity cannot be confirmed unless data for all four of these cells are gathered. Such Bayesian thinking seems essential to improving clinical judgment (see also Arkes, 1981).

The second test is to make sure that the sign increases diagnostic accuracy. A sign may be strongly associated with a disorder, but it also must not falsely identify pathology at a rate greater than the frequency of the pathology itself. Finally, one must assure oneself that use of a diagnostic sign results in incremental validity. That is, the sign must not be redundant with other diagnostic signs; it must contribute new information.
or predictive power. If it does not, the clinician can become overconfident being falsely reassured by each new (but redundant) sign. It is rare when more than a handful of diagnostic signs are needed to maximize predictive accuracy.

**Using Disconfirming Strategies**

Faust (1986) also noted that clinicians would be more accurate if they emphasized the disconfirmation of hypotheses rather than confirming evidence for an a priori hypothesis. Unless one accurately searches for disconfirming information, one often will not obtain evidence that a hypothesis is absolutely wrong. “Signs associated with more than one condition or with many conditions (e.g., anxiety) are much less useful for differential diagnosis than exclusionary signs or criteria” (p. 427). As a general rule, the clinician should try to generate several alternative explanations for their clinical hypotheses. In particular, it is helpful to generate a set of confirmatory (“why might I be correct?”) and a set of disconfirmatory (“why might I be incorrect?”) explanations. Multiple competing hypotheses should be considered so that the data obtained can rule out some while supporting an alternative.

When individuals are forced to generate alternative explanations for outcomes, errors in reasoning such as the hindsight bias are reduced because generated alternatives are then judged as more likely to occur (Ross, Lepper, Strack, & Steinmetz, 1977; Slovic & Fischhoff, 1977), thus reducing overconfidence and consequently premature closure (Elstein, Shulman, & Sprafka, 1978). Generating multiple alternative explanations also increases the probability that one of them might be correct (Arnoult & Anderson, 1988).

**Beware of Client Acceptance of Faulty Judgment**

The term *Barnum effect* has been used by Meehl (1956) to describe the situation in which people willingly accept personality interpretations from experts even though they are based on vague horoscopelike statements that are likely to apply to most people in the general population (e.g., “Although you generally think well of yourself, at times you doubt your abilities”). This remarkably consistent phenomenon was reviewed by Snyder, Shenkel, and Lowery (1977), who noted that acceptance of Barnum interpretations is enhanced by situational features common to the clinical diagnostic situation. The client assumes that the interpretation was specifically developed for him or her, derived from the results of valid psychological assessment techniques, and delivered by an expert, high-status clinician. After accepting such interpretations, clients increase their faith in psychological testing and their confidence in the skills of their clinician (Snyder, Larsen, & Bloom, 1976; Snyder & Shenkel, 1976).

It is important that clinicians not be reinforced by their clients’ enthusiastic acceptance of personality interpretations (even when they are not Barnum statements). The Barnum effect literature establishes clearly that clients will accept any reasonable-sounding interpretation delivered by an authority figure so long as it seems somewhat tailored to them personally. We have even been able to replicate the Barnum effect in our Abnormal Psychology classes in which all students take a brief personality scale and later in the semester all receive the same description of their personalities on a form with their personalized code number. Nearly all of these students rate the description as a good or excellent formulation of their personality characteristics. Snyder et al. (1977) aptly pointed out that the “sucker” in the Barnum situation may not be the naive client who believes generic feedback, but the clinician who interprets such client acceptance as valid feedback about his or her clinical skills.

**Recognizing Predictive Uncertainty**

Given that prediction is filled with uncertainty, clinicians should accept this uncertainty and recognize that a certain amount of error is avoidable. By accepting such fallibility, clinicians will not too hastily abandon good but imperfect predictors and will not be overconfident in stating predictions and diagnoses. Recognition that as clinicians we are fallible should allow us to accept research on judgment and decision-making processes less defensively.

**Minimizing Reliance on Memory**

As discussed earlier, most judgmental biases are the result of limitations of our cognitive systems to deal simultaneously with multiple bits of information. To minimize bias, then, we should be motivated to minimize reliance on memory. Along these lines, Dawes (1986) suggested that it is helpful to rely on external aids—the computer, pencil and paper—when we need to estimate frequencies, and that it makes sense to actually write down base-rates and probability ratios. Unfortu-
nately, decision-makers often reject such external aids. They often believe that they can improve on them by using inferences based on theoretical viewpoint and clinical experience, even though these are likely to worsen prediction (Ark es, Dawes, & Christensen, 1986).

Minimizing reliance on memory has other useful functions. For example, it can reduce the likelihood that information not presented at a case conference that is consistent with a case summary is misremembered as having been presented. Unpresented symptoms that are consistent with a diagnosis tend to be remembered as having been presented (Ark es & Harkness, 1980). Moreover, previously presented symptoms that are not consistent with the diagnosis are often forgotten.

There are a multitude of creative solutions to reducing the load on memory during judgmental tasks. For example, Arnould and Anderson (1988) suggested using multiple judges, role-playing, videotape, rating scales, balance sheets, graphics, and hypothetical questions. And Nisbett and Ross (1980) have provided some helpful maxims. By not burdening working memory, clinicians may be less likely to rush to choose a seemingly satisfactory but perhaps wrong solution to a judgmental problem.

**IMPROVING RESEARCH ON CLINICAL JUDGMENT**

After nearly 40 years of research on judgment, inference, and decision-making, the influence of this corpus of data on actual clinical practice is not very evident. Few volumes have been written attempting to translate research of laboratory-oriented social and cognitive psychologists into a language understandable by the practicing clinician (but see Turk & Salovey, 1988). The investigators complain that clinicians react defensively to their suggestions concerning improving judgmental accuracy. But often these suggestions have been made in the context of books and articles ignoring or discounting the difficulties faced by the clinician and, at times, questioning the intelligence of the clinician him or herself. For research on judgmental processes to have any impact on clinical practice, such “clinician bashing” must cease. Clinicians may operate from a different “assumptive world” than decision theorists, but if presented empathically, the messages of such investigators need not be lost on them. To the extent that psychology as a discipline can reaffirm its commitment to training clinicians who are also scientists (the fading scientist-practitioner model), we increase the likelihood that important research on judgment and decision-making will find its way to the clinic.

Recently, the clinical judgment literature has also been criticized as suffering from an inadequate framework for understanding different types of judgment and, especially, for failing to attend to the context in which judgments are made (Rock, Bransford, Maisto, & Morey, 1987). These authors suggest that a more “ecological” approach be taken in studying judgment that explicitly attends to “(a) characteristics of the therapist, (b) information processing strategies made available to the clinician, (c) critical tasks that define the major focus of specific judgments, and (d) the nature of the clinical materials that provide the basis for judgments” (p. 645).

For many years, Holt (1958, 1961, 1970, 1988) and others have argued that the literature on judgment, which is based primarily on laboratory experiments, may not be directly generalizable to the kinds of clinical judgment tasks faced by the practicing clinician. He has described six discrepancies between clinical judgment as studied versus practiced. As summarized by Rock et al. (1987), Holt questioned the ecological validity of the extant research on clinical judgment. The debate about ecological validity of the laboratory paradigm in cognitive psychology now rages in a variety of contexts (compare Neisser, 1982, with Banaji & Crowder, 1989). It is sufficient here to note, however, that Rock et al. (1987) suggested that the experimental literature on clinical judgment would have greater impact if relevant variables affecting the context in which judgments are made are considered. For example, they suggested (based on Jenkins, 1979) (a) that investigators measure the explicit characteristics of their subjects relevant to the judgment task such as abilities, knowledge, preferences, experience, and orientation, and (b) that the criterial tasks—diagnosis, prognosis, formulation, and description—be made explicit. Furthermore, they recommended that characteristics of the clinical materials on which judgments are based be attended to, such as in-person interviews, taped interviews, case notes, case conference, case history, test battery, or the reports of others; and that the specific information-processing activities required of the clinician be measured explicitly: single versus multiple judgment trials, opportunity for
feedback, and ability to debug problems in making judgments.

In analyzing the data collected from experiments considering all of these variables, interactions should be emphasized. What combinations of variables accumulated in what ways produce especially good or especially poor judgmental outcomes? In recent years, the ecological approach to clinical judgment has motivated reviews of the situational factors impinging on person perception during the clinical interview (Cline, 1985) and studies of the judgments of physicians (LaDuca, Engel, & Chovan, 1988).

**SUMMARY**

This chapter describes some of the difficulties faced by clinicians in the course of making judgments at five stages in the diagnostic process: (a) forming a first impression, (b) using a classification system, (c) making a diagnosis, (d) assessing covariation, and (e) making decisions and predicting the future. At each of these steps, various factors conspire to produce suboptimal judgment mostly emanating from the limited capacity of our cognitive systems. Several suggestions are made concerning the reduction of these biases, and we conclude with some thoughts regarding how to make research in this area more accessible to the practicing clinician in his or her daily activities. In particular, clinical judgment is improved when diagnosticians examine the validity of tests very carefully, use a disconfirming rather than confirmatory hypothesis-testing process, reduce reliance on client acceptance of interpretations as evidence of their validity, accept the presence of predictive uncertainty, and use any expedient possible to reduce demand on working memory. Conversely, research on clinical judgment will be more likely to influence the practicing clinician when it attends to the real conditions in which clinical decisions are made and is communicated in a context that acknowledges the difficulties inherent in clinicians' work. Perhaps the scientist-practitioner perspective that is being revived through the health psychology movement will serve to increase communication between investigators of clinical judgment and clinicians themselves.

**REFERENCES**


Langer, E. J. (1975). The illusion of control. Jour-
nal of Personality and Social Psychology, 32, 311–328.


Snyder, M., & Cantor, N. (1979). Testing hypotheses about other people: The use of histor-


Weiner, B. (1975). "On being sane in insane
places:” A process (attributational) analysis and critique. *Journal of Abnormal Psychology, 84*, 433–441.


