1994

The Base-Rate Fallacy in School Psychology: Implications for Decision-Making

Mary Lynne Kennedy
University of Rhode Island

THE BASE-RATE FALLACY IN SCHOOL PSYCHOLOGY:

IMPLICATIONS FOR DECISION-MAKING

BY

MARY LYNNE KENNEDY

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND
1994

Abstract

This dissertation generalized the base-rate fallacy in diagnostic decision-making to the field of school psychology. This fallacy involves allowing information other than prior probabilities inappropriately to dominate probability judgments. Participants included 80 school psychologists who were asked to rate the probability of a learning-disability diagnosis and to state their degree of confidence in that diagnosis. In Study I, participants received only base-rate information. In Study II, participants received either (a) relevant base-rate information plus irrelevant individuating information, (b) relevant base-rate information plus relevant diagnostic information, or (c) relevant base-rate information plus relevant diagnostic information plus irrelevant individuating information. Additionally, in Study II these three levels of information were completely crossed with a second independent variable, that is, salient link. Here, base-rate information was either linked or not linked with the diagnosis to be predicted. Results showed that (a) school psychologists appropriately used relevant base-rate information only when no other information was available, (b) when a salient link between relevant base-rate information and the diagnosis to be predicted was provided, diagnostic accuracy did not change, (c) when relevant diagnostic and irrelevant individuating information were provided in addition to base-rate information, diagnostic accuracy decreased relative to base-rate-only information, and (d) school psychologists were least confident of diagnoses for which they demonstrated most accuracy. Results were discussed in terms of implications for diagnostic decision making in the field of school psychology.

Acknowledgment

There are several people who provided ongoing help and assistance in completing this dissertation, but first and foremost, I would like to thank my major professor, Dr. W. Grant Willis. His insight, direction and clarity of thought was invaluable. There are no words to adequately express my appreciation for all he has contributed to this dissertation, as well as to my own professional development. I was extremely fortunate to have had Dr. David Faust as a member of my dissertation committee. I owe him a profound debt of gratitude for his ongoing support, advice, and insight in developing and completing this dissertation. I also want to thank Dr. Susan Brady and Dr. Dayle Joseph who helped to develop and refine this dissertation with their unique perspectives and ideas. I do not know how to begin to thank Dr. Michael Weiler who continually and unselfishly put aside his own work to answer "just one more question". His calm demeanor and sense of humor helped me make it through the most insane moments. He has been a constant source of inspiration and I feel very lucky to have him as a part of my life. Finally, I want to acknowledge my mother who deserves a part of this PhD. Without her, none of this would have been possible. I am very fortunate to have such a loving and unselfish role model.

Table of Contents

List of Tables

Introduction

This dissertation examines the use of base rates by school psychologists in a diagnostic decision-making task. Current knowledge of diagnostic decision-making in school psychology is important because if suboptimal reasoning is used to make decisions about children and youth with disabilities, it is likely that many students will be misclassified. For example, children in need of special education and related services may not be identified, whereas those who do not need such services may be labeled as disabled. Children mislabeled as in need of services may then suffer social and emotional consequences. Fagley (1988) has called for research that examines whether school psychologists are prone to the same decision-making errors (i.e., underutilization or neglect of base rates) as has been documented in other fields (see Eddy, 1982; Schwartz, Gory, Kassirer, & Essig, 1973).

The present study was designed to investigate psychoeducational decision-making with respect to the processing of base-rate information and to examine conditions that may increase diagnostic decision-making accuracy. Base rates refer to the frequency of an event or diagnosis and are integral to diagnostic decision-making. Bayes theorem provides an explicit model of how to incorporate base-rate information into the decision-making process; however, research in the area of decision-making has documented that both clinicians and laypeople are subject to a systematic and predictable decision-making error, that is, the "base-rate fallacy" (Bar-Hillel & Fischhoff, 1981; Kahneman & Tversky, 1973; Lyon & Slovic; 1976). This fallacy involves allowing information other than base rates to dominate when making probability judgments. Thus, base-rate information often is overlooked even though it is essential to probabilistic decision-making.

In the following section, several areas of literature are reviewed that provide the theoretical underpinnings for this study. These include: (a) the research within cognitive psychology that examines the cognitive processes of human judgment and decision making, (b) the cognitive domain of social psychology that examines the structure and use of knowledge and judgment rules, and (c) the research on base rates that documents their role in decision-making and the effects of their neglect.

<u>Research in Cognitive Psychology</u>

Research in human judgment and decision-making has several points of origin. Cognitive psychology's focus on human judgment and decision-making stemmed from research conducted in the 1940s and 1950s. This research, which compared statistically derived judgment models, or actuarial models, to the clinical decision-making processes of the human judge, revealed that the actuarial approaches outperformed the clinical approaches (Meehl, 1954). Subsequently, cognitive psychologists began to focus on understanding the cognitive processes of the human judge.

The information-processing approach within cognitive psychology used the computer to model the hidden decision-making strategies of the human judge. One of the main tenets of information-processing theory is that the human information-processing system is limited due to a focused window of attention, as well as a limited short-term memory capacity (Miller, 1956). These cognitive limitations have obvious implications for decision making.

Kahneman and Tversky (1973), in their seminal work, identified decision-making errors characteristic of human judges due to the limitations of the human information-processing system. They studied the coding and combining of information by human judges in probabilistic decision-making tasks. Their research consistently documented a discrepancy between judges actual performance and optimal performance. This suboptimal performance has been attributed to the use of cognitive heuristics or judgment rules that reduce complex tasks, such as assessing probabilities and predicting values, to simpler judgmental operations.

These heuristics are considered to be an important feature of our cognitive systems because of the limitations of human information processing (Tversky &

Kahneman, 1974) and because of the complexity and the variety of decisions we make (Nisbett & Ross, 1980). It has been documented, however, that the same processes and tools that can simplify a problem, when overapplied, also can lead to systematic and predictable errors (Tversky & Kahneman, 1974).

Thus, Tversky and Kahneman (1974) focused on gaining a better understanding of human information processing by examining: (a) the cognitive heuristics used to reduce complex decision-making tasks, and (b) the kinds of errors that are made through the use of these heuristics. They explained the justification for this research strategy:

> There are three related reasons for the focus on systematic errors and inferential biases in the study of reasoning. First, they expose some of our intellectual limitations and suggest ways to improve the quality of our thinking. Second, errors and biases often reveal the psychological processes that govern judgment and inference. Third, mistakes and fallacies help the mapping of human intuitions by indicating which principles of statistics or logic are non-intuitive or counter-intuitive. (Kahneman & Tversky, 1982; p.124).

Heuristics. There are three main cognitive heuristics that can lead to systematic decision-making errors in prediction, in part, because they are not sensitive to base-rate information. First, the representative heuristic refers to a decision-making strategy that relies on making a decision or prediction based on similarity. For example, research examining the representativeness heuristic has determined that people assess probabilities by the degree to which the item to be predicted is similar or representative of their conception of a category (Kahneman & Tversky, 1972). One of the main problems with making probability judgments by representativeness is the neglect of prior probabilities, or base rates. Base rates have no effect on the representative heuristic, even though they should play a major role in assessing probabilities.

In other situations, people make probability judgments based on the availability or ease with which the category to be predicted can be brought to mind. The availability heuristic can be a useful tool for assessing probability because in many cases larger classes of events are retrieved more quickly; however, availability also is influenced by other factors that have no bearing on the probability of occurrence. Specifically, availability can be influenced by familiarity, salience, the effectiveness of a search set, illusory correlation, and imaginability (Tversky & Kahneman, 1974). The influence of these factors leads to neglect of the base rates, and thus to errors in probability judgments.

Anchoring and adjustment, the third heuristic, rely on making probability judgments by starting from an initial value (anchor) and adjusting that value to reach a final estimate. This strategy of anchoring and adjustment has been demonstrated to result in predictable errors in which people insufficiently adjust their estimates by anchoring on the initial value (Tversky & Kahneman, 1974). Anchoring can result when an initial starting value is provided or when an incomplete computation is made of the available data (e.g., failure to incorporate base-rate information).

Summary. In summary, the research within cognitive psychology has attempted to clarify the underlying cognitive processes of the human judge. The results of this research have documented the limitations of our information-processing systems and the resultant suboptimal reasoning strategies. Cognitive heuristics are a means by which we simplify complex problems, such as assessing probabilities. These heuristics, however, can lead to serious judgment errors, in part, because they neglect base-rate information. Thus, one of the most important approaches used in studying judgment and decision making has been to analyze the systematic discrepancies between normative judgment models and actual judgments.

Research in Social Psychology

A cognitive approach to the study of human judgment within social psychology, which closely paralleled the study of human judgment within cognitive psychology, began to gain attention in the 1960s. This cognitive orientation stemmed from research such as Tajfel's (1969) "Cognitive Aspects of Prejudice" that applied a cognitive perspective to the understanding of human behavior. In the early 1970s, the longstanding attitude--behavior link primarily used to explain human judgment was criticized, and a new emphasis was placed on the information-processing limitations, not the motivations, of the human judge (Israel & Tajfel, 1972). Further, the use of the computer to model human judgment was de-emphasized within cognitive psychology; instead, understanding the limitations of human judgment was stressed.

Schemas. Researchers in social psychology, as well as in cognitive psychology, emphasized the role of schemas in aiding human judgment and information processing. Schemas are cognitive structures that are formed by past experiences that help us to organize and to understand information (Neisser, 1976). They provide meaning and organization to new events and objects, and help to direct our attention to relevant features of the environment. Thus, they provide a structure by which new material can be integrated with memories that have been stored. Based on this role, schemas can aid us in inferring and predicting what will happen.

Schemas, like heuristics, however, can lead to biases and errors in judgment. These errors can occur when we miscategorize information or we oversimplify our processing of information. For example, our schemata can sometimes misdirect our attention and, thus, we ignore relevant information. We also can make errors by remembering schema-related details that did not actually occur (Schneider, 1988). Heuristics also may guide the selection of schemas and, thus, inappropriate schemas may be activated (Nisbett & Ross, 1980). Further, it has been suggested that humans lack good schemata for dealing with probabilistic data. It appears that although the logical importance of base rates cannot be overstated, this kind of information fails to be effective in activating appropriate schemas.

Stereotypes can be thought of as schema-like concepts in that they guide us in encoding, retaining, and retrieving information in the social world. Hamilton and Rose (1980) suggested that we can understand stereotypes better by framing them as correlational concepts. For example, the stereotype that women are bad drivers, is based on a perceived relationship between an identified group, that is, women, and some trait, that is, bad driving. If we examine how people form concepts and beliefs based on the covariation between two variables, we may understand stereotypes better.

Covariation analysis. The cognitive demands involved in accurately assessing covariation make us prone both to the development and maintenance of stereotypes (Hamilton & Rose, 1980). Determining when two (or more) variables truly are related, that is, covariation analysis, is a difficult cognitive task. In order to establish that a relationship exists between two variables, in this case a particular group and some trait, one must show that the trait occurs more frequently among that group than among other groups. Individuals have great difficulty analyzing covariation and thus commonly draw false conclusions about relationships between two variables (Arkes & Harkness, 1980; Gnys, Willis, & Faust, in press; Smedslund, 1963; Ward & Jenkins, 1965).

With respect to diagnostic decision making, this error occurs when decision makers consider only the frequency with which the sign is present and the disorder present, and fail to consider other situations in which (a) the sign is present but the condition absent, (b) the sign is absent but the condition is present, and (c) the sign is absent and the condition is absent. Failure to consider all of this information makes it impossible to determine whether a valid relationship exists between the sign and the condition.

These false associations between variables have been referred to as illusory correlations (Chapman & Chapman, 1967). The disregard of base rates is one of the main contributing factors to the formation of illusory correlations. Here, the psychologist comes to believe that particular features of test performance provide trustworthy indicators of certain diagnostic categories, without attending to the base rates (Chapman & Chapman, 1967). For example, one commonly reported relationship interpreted from the Draw-a-Person Test is that highly detailed drawings of eyes are associated with paranoia. Although research has failed to validate this (and many other sign-diagnosis relationships) on the Draw-a-Person Test, clinicians still often agree about such clinical correlates of Draw-a-Person performance (Ziskin & Faust, 1991).

Conclusions from Cognitive and Social Psychology

This section has highlighted the research in both cognitive and social psychology on human judgment and decision-making. Research has revealed some of the limitations of the human judge, and the cognitive strategies and rules used to compensate for these limitations. For example, cognitive heuristics and schemas facilitate processing of complex data. These decision aids, however, can lead to systematic errors, especially when dealing with probabilistic data, in part because they neglect base-rate information. The neglect of base rates has been implicated in diagnostic decision-making errors, as well as in errors of judgment in the social world. For example, stereotypes can be thought of as illusory correlations that develop and are maintained, in part, due to a neglect of base-rate information. In the following section, base rates and their role in diagnostic prediction are reviewed and the potential implications of the base-rate fallacy for the field of school psychology are discussed.

Base Rates

Base rates, as stated previously, refer to the frequency with which an event (i.e., a symptom or condition) occurs. For example, if school phobia exists in 1 out of 100 junior high-school students, the base rate for school phobia in that population is 1%. Thus, base rates refer to the prevalence of a characteristic in a population.

One of the problems associated with the neglect of base-rate information, as discussed in the section on formation of stereotypes, is that a clinician may not know whether a sign is validly related to a condition. Knowledge of whether a sign is validly related to a condition, however, is not enough to determine whether use of that sign will be helpful in diagnostic prediction (Faust & Nurcombe, 1989). The clinician still needs to know the base rate or frequency of the condition in the population. If one lacks knowledge of the base rate of the condition, then it is impossible to determine the diagnostic utility of the sign, that is, whether more erroneous classifications will be made when using the sign. The validity of diagnostic or predictive signs is never absolute or constant, but must be interpreted in relation to the prior odds or the base rates (Meehl & Rosen, 1955). Disorders that occur less frequently require a more accurate sign if more correct than false identifications are to be made. If use of a sign does not improve the accuracy one would achieve by relying on the base rates alone, then overall accuracy will be decreased by relying on the sign (Faust & Nurcombe, 1989).

The following example is adapted from Faust and Nurcombe (1989). Suppose that the prevalence of school phobia is 1 out of 1000 cases. School psychologist A relies on the base rates and consistently says that the condition is not present. School psychologist B only diagnoses the condition when there is overwhelming evidence. Further, suppose that school psychologist B never misses a true case of school phobia, and only 1 out of 100 times, says the condition is present when it is not. On a sample of 1000 cases, school psychologist A will make one error, that is, will miss one case of school phobia when it is present. School psychologist B, on the other hand, will make approximately ten errors, that is, will detect the one case of school phobia, but also will misidentify nearly ten children (i.e., almost 1 of 100).

Thus, base rates are integral to diagnostic decision making both to determine: (a) whether a sign is validly related to a condition and, further, (b) whether use of that sign will increase diagnostic accuracy. As Faust and Nurcombe (1989) suggested, however, base rates should not be the sole criteria by which we guide our practice. Diagnostic decision making should be framed as a analysis of the kinds of errors we are willing to make (e.g., the consequences of overdiagnosing a condition versus missing a condition). In some situations, missing a condition may have relatively minor consequences; in other circumstances, missing a condition can have life threatening consequences.

Ignoring base rates. Research examining base rates in probabilistic decision making tasks has typically pitted base-rate information against some kind of individuating or diagnostic information (e.g., a personality description or the correct positive rate and false-alarm rate of a test) (Arkes, 1989). A consistent finding of this research is that individuals presented with base-rate information and case-specific, or individuating information, will defer to the individuating information (Kahneman & Tversky, 1973). For example, in their seminal work, Kahneman and Tversky (1973) presented participants with base-rate information (e.g., percentage of people employed as lawyers versus engineers), and individuating information (e.g., a brief personality description). The personality description was designed to be non-diagnostic in differentiating lawyers from engineers. Participants were asked to make a probability judgment (on a scale of 0 to 100) as to whether the person was an engineer or a lawyer. Results of their study demonstrated that participants relied on the stereotypic description and ignored the base-rate data. Participants did use base-rate data when no other information was presented.

A similar finding was documented by Borgida and Nisbett (1977). In their study, participants were presented with base-rate information in the form of mean course ratings, and individuating information in the form of former students' descriptions of the course. Again, participants relied on the individuating information for course selection and neglected the base rates or mean course ratings.

Probability judgments also can be dominated by an accuracy effect (i.e., anchoring on the diagnosticity of specific information) (Lyon & Slovic, 1975). Tversky and Kahneman (1980) conducted a study in which participants were presented with a scenario and asked to make a probability judgment. They were told that 85% of the cabs in the city are blue and 15% are green. A cab has been involved in an accident and a witness has identified the cab as green. The witness was accurate in identifying each color 80% of the time and confused colors 20% of the time. Participants were asked to estimate the probability that the cab involved the accident was green as the witness had stated. In this Bayesian-inference problem, all normatively relevant information (i.e., base rates, correct positive rate, and false positive rate) was supplied.

Bayes rule provides a normative formula for this decision:

$$P_j = Pp_1 / (Pp_1 + Qp_2) \quad (1)$$

Here, P_j is the probability that the cab is green; P is the base rate of green cabs in the city, p_1 is the valid positive identification rate, Q is the base rate of blue cabs in the city, and p_2 is the false positive identification rate. Applying Equation 1 to the previous example, the correct probability judgment is 41%, that is, .41 = (.15 x .80) / [(.15 x .80) + (.85 x .20)]. Instead, participants tended to anchor on the hit-rate information without integrating pertinent base-rate information, and made estimates of 80%.

Studies have been conducted examining the stability of the base-rate fallacy (Bar-Hillel, 1975; Lyon & Slovic, 1976). For example, variations of one of Tversky and Kahneman's original base-rate problems were presented to 350 participants. Some of the variables manipulated included order of information presentation, varying the extremity of the base rates (i.e., 90% versus 60%), and using a verbal description instead of actual statistics for the base-rate data (i.e., using words such as most instead of actual numbers). Results of this research consistently demonstrated that when other information was present, the base-rate information was neglected. Thus, the base-rate fallacy is considered to be a robust phenomenon (Ajzen, 1977; Bar-Hillel, 1980; Lyon & Slovic, 1976; Nisbett & Borgida, 1975).

Causality. In an attempt to encourage people to use base-rate information, cognitive psychologists have attempted a variety of interventions (Arkes, 1989). One approach has been to enhance the causal link between the base-rate information and the category to be predicted. Nisbett and Ross (1980) concluded that individuals may lack an intuitive understanding of Bayesian statistics, but their judgments would be much less impaired if base rates could be made more salient. Situations in which base rates appear more salient include conditions where no other information is presented or where the base rate has a causal interpretation.

Studies have manipulated causal relevance and demonstrated that base rates may dominate probability judgments in particular conditions (Ajzen, 1977). For example, in a study conducted by Ajzen, some participants were given a brief description of a student, Gary. Participants then were asked to make a probability judgment as to whether Gary passed or failed an exam. The first group of participants was told that Gary had been randomly sampled from a group of students, among which 75% had failed the exam and 25% had passed the exam. These participants ignored the normative relevance of the base rates in favor of the individuating information. Another group of participants, however, was told that the test was failed by 75% of the people in Gary's class and passed by 25% of the people in Gary's class. These individuals did attend to the high base rate of failures (i.e., 75%) and did not defer to the individuating information. Thus, Ajzen concluded that base rates that are causally linked to the event or target to be predicted may influence probability judgments.

Bar-Hillel (1980) discussed the relationship of causality to relevance, proffering that causality is only one aspect of relevance. She suggested that individuals may ignore base-rate information because they think it is irrelevant. Moreover, she stated that if the relevance of base-rate information is enhanced, individuals may be more likely to use this information. Results of her study indicated that two items of information may be used only if they appear equally relevant. Thus, one way relevance can be manipulated is by providing a causal link between the base-rate data and the outcome to be predicted. In this way, the base rate is interpreted as having some causal relation to the outcome to be predicted.

Research Questions

This dissertation was an investigation of whether the base-rate fallacy generalized to the field of school psychology. Further, conditions were varied to determine whether there are conditions favorable to the use of base-rate information. Specifically, this study examined the effects of illusory-correlation information, accuracy-rate information, and salience of link, on the use of base-rate information. To clarify terminology, accuracy-rate information has been referred to in the research literature as diagnostic or true-positive rate information, illusory correlation has been referred to as individuating information or a belief in a relationship that does not exist, and salience of link has been referred to as relevance or causality.

In this study, the illusory-correlation was related to the diagnosis of a learning disability. The illusory-correlation information was included to elucidate how this kind of information, in combination with base-rate information and accuracy-rate information affects diagnostic accuracy. This is considered to be an important area of investigation as research has demonstrated that school psychologists are prone to the phenomenon of illusory correlation.

The choice of the illusory correlation was based on research (Gnys et al., in press) that has documented that some school psychologists believe in a systematic relationship between levels of intersubtest scatter on the Wechsler Intelligence Scale for Children--Revised (WISC-R) and the diagnosis of learning disabilities, although research has failed to validate this relationship (Kaufman, 1979). Moreover, research has documented that normal children show a mean range of seven (SD = 2) scaled points on the WISC-R (Kaufman, 1976). Thus, a scaled-score difference as large as nine points is unremarkable (Reynolds & Kaufman, 1990). These findings document that intersubtest scatter has negligible predictive value in the diagnosis of learning disabilities. Further, warnings have been advanced (Kaufman, 1976) that levels of scatter that frequently occur in the normal population should not influence diagnostic decisions.

The diagnostic information provided the accuracy rate of the test (i.e., the percentage of children diagnosed as learning disabled who are learning disabled) and the false-positive rate (the number of children identified as learning disabled who are not learning disabled). In conditions where accuracy-rate information was presented, all the normatively relevant information necessary to make a Bayesian calculation was presented. Finally, salience of link was explored to examine whether providing a link between the base-rate information and the category to be predicted would increase diagnostic accuracy. These variables were manipulated systematically to provide directions for future intervention considerations.

In conclusion, this dissertation addressed the following questions:

1. Do school psychologists use base-rate information? It was predicted that school psychologists would use base-rate data when no other information was presented.

2. Does providing a salient link between the base rate and the category to be predicted increase accuracy? It was predicted that school psychologists' probability judgments would be more accurate when there was a salient link between the base rate-information and the category to be predicted.

3. Does the kind of information presented with the base-rate information affect accuracy? It was predicted that school psychologists would be less accurate when presented with base-rate information plus illusory-correlation information plus accuracy-rate information versus when presented with either (a) base-rate information plus accuracy-rate information or (b) base-rate information plus illusory-correlation information..

The first prediction was based on research showing that when base-rate information is presented alone, both lay people and professionals use that information appropriately when making probability judgments (e.g., Kahneman & Tversky, 1973). The second prediction was based on limited research showing that when a salient link between the base-rate information and the category to be predicted is provided, judgmental accuracy improves (Ajzen, 1977). Finally, the third prediction was based on two areas of research showing that judgmental accuracy decreased when base-rate information was combined either (a) with relevant diagnostic information (e.g., Lyon & Slovic, 1975) or (b) with irrelevant individuating information (e.g., Kahneman & Tversky, 1973). Thus, it was predicted that lower levels of judgmental accuracy would result when base-rate information is combined with both relevant diagnostic information and irrelevant individuating information, than when base-rate information is combined with either of these other sources of information alone.

Method

Study I

Subjects. A preliminary, descriptive study was conducted to determine if school psychologists would, indeed, use base-rate information when no other information was provided. Information packets were mailed to a random sample of 60 school psychologists (practitioners and trainers) selected from the Directory of Nationally Certified School Psychologists (NASP, 1991). Psychologists who did not respond within four weeks were sent a second mailing. All participants were guaranteed anonymity and were treated in accordance with ethical standards adopted by the American Psychological Association (APA, 1981) and by the University of Rhode Island Institutional Review Board (1992).

Materials and procedure. All psychologists received a cover letter that explained the purpose of the survey and guaranteed their anonymity. A copy of this letter appears in Appendix A. The information packet comprised two parts. In the first section, respondents indicated: (a) highest academic degree, (b) year of degree, (c) gender, (d) geographic location, and (e) primary professional role (e.g., practitioner, trainer). In the second section, the packet included a case scenario providing base-rate information, specifically the percentage of boys diagnosed as learning disabled in a school (i.e., 10%). The case scenario used in Study I appears in Appendix B.

Respondents first filled out the demographic information on the questionnaire. They then were asked to rate the probability that the child described by the case scenario was learning disabled on a scale from 0% (lowest--not probable at all) to 100% (highest--certain). They also were asked to rate their confidence in their decision on a Likert-type scale ranging from 1 (not confident) to 7 (very confident). Respondents' probability ratings were used to calculate an accuracy score. This score represented the absolute difference between the participant's stated probability rating and the correct rating of 10%. Thus, lower difference scores represented higher accuracy.

Study II

Subjects. Information packages were mailed to a random sample of 300 school psychologists (practitioners and trainers) selected from the Directory of Nationally Certified School Psychologists (NASP, 1991). Student affiliates were excluded from the sample. Psychologists who did not respond within approximately four weeks received a second mailing in order to obtain at least 20 completed surveys in each condition.

Materials and procedure. All psychologists received a cover letter that explained the purpose of the survey and guaranteed their anonymity (see Appendix A). Materials consisted of six different information packets, each comprising two parts. Similar to Study I, in the first section, respondents indicated: (a) highest academic degree, (b) year of degree, (c) gender, (d) geographic location, and (e) primary professional role (e.g., practitioner, trainer).

The second section contained case-scenario information. Three kinds of information were provided in varying combinations. This included (a) base-rate information; (b) irrelevant, individuating information designed to assess susceptibility to illusory correlation; and (c) relevant, diagnostic accuracy-rate information. Every packet included base-rate information, specifically the percentage of boys diagnosed as learning disabled in a school (i.e., 10%). Two-thirds of the packets contained the irrelevant information. This was provided by including the summary page of the response booklet from the WISC-R. This summary page showed a range of intersubtest scatter of 9 scaled-score points. This range was chosen because of its fairly common occurrence and its lack of diagnostic utility (Kaufman, 1979; McDermott, Gluttin, Jones, Watkins, & Kush, 1989; Reynolds & Kaufman, 1989). All IQ scores were within the average range (Verbal IQ = 97, Performance IQ = 104, and Full Scale IQ = 100). Two-thirds of the packets included the relevant accuracy-rate information pertaining to a procedure used to diagnose learning disabilities. This included the true-positive rate of the procedure (i.e., 80%) and the false-positive rate (i.e., 20%). Case scenarios such as these have been used in research examining clinical decision making and have been demonstrated to be highly correlated with actual cases (Kiriwan, Chaput de Saintonge, Joyce, & Currey, 1983).

Finally, in addition to the case-scenario information, half of the packets contained a statement clearly linking the base-rate information with the diagnosis to be predicted (i.e., "Because of an effective early-intervention program in John Smith's school, the incidence of learning disabilities among boys in this school is only 10%."), whereas the other half did not contain this statement. The case scenarios used in Study II appear in Appendix C.

Respondents first filled out the demographic information on the questionnaire. They then were asked to rate the probability that the child represented by the case-scenario information was learning disabled on a scale from 0% (lowest--not probable at all) to 100% (highest--certain). They also were asked to rate their confidence in their decision on a Likert-type scale ranging from 1 (not confident) to 7 (very confident).

Design. Each participating school psychologist was assigned randomly to one of six conditions created by varying the levels of two independent variables. One independent variable, salience of link (i.e., provision of a link between the base-rate data and the category to be predicted), comprised two levels: (a) salient link between the base-rate data and category to be predicted and (b) no link provided between the base-rate data and category to be predicted. The other independent variable, information, comprised three levels: (a) base-rate information plus relevant, that is, accuracy-rate, information (BR + AR), (b) base-rate information plus irrelevant, that is, illusory-correlation, information (BR + IC), and (c) base-rate information plus accuracy-rate information plus illusory-correlation information (BR + AR + IC).

The levels of these variables were completely crossed, resulting in a two-by-three between-groups design. Specifically, one group received a case scenario defined by BR + AR information with no link, one by BR + AR information with a salient link, one by BR + IC with no link, one by BR + IC with a salient link, one by BR + AR + IC with no link, and one by BR + AR + IC with a salient link.

Similar to Study I, the dependent variables were (a) a diagnostic accuracy rating, and (b) a confidence rating in the diagnosis. Accuracy ratings were derived by calculating the absolute difference between the participant's stated probability judgment and the correct value for the materials received. The correct value for the BR + IC condition was 10% (the base rate), whereas the correct value for the BR + AC and the BR + AR + IC conditions (as calculated from Equation 1) was 31%.

Results

Analysis I

Of the 60 surveys mailed for Study I, 13 were returned, and upon a remailing of the remaining 47, 9 were returned. This yielded a final response rate of 37%. Two surveys were deleted randomly in order to obtain 20 surveys for the group. The sample characteristics for Study I are presented in Table 1.

Table 1

Sample Characteristics for Study I

Characteristic	n	%
Educational Level		
Master's	3	15
Master's + 30	14	70
Doctorate	2	10
Other	1	5
Region		
North East	3	15
North Central	6	30
West	4	20
South	7	35
Primary Role		
School Psychologist	17	85
Administrator/Supervisor	0	0
University Trainer	2	10
Other	1	5
Year of Terminal Degree		
1988-1993	6	30
1983-1987	4	20
1978-1982	5	25
1973-1977	3	15
1972 or before	2	10

Results of Study I revealed that all respondents used base-rate information accurately when no other information was presented (M = 0, SD = 0). That is, there was no difference between the reported and the optimal probability judgment for any of these respondents. The mean confidence rating was 2.7 (SD = 1.7).

Analysis II

Of the 300 surveys mailed for Study II, 106 were returned, yielding an initial return rate of 35%. Remailings were sent to the 194 nonresponders; 28 were returned yielding a remailing response rate of 14%. This procedure resulted in a total of 134 surveys, or a response rate of 45%. The number of surveys in each group ranged from 20 through 27. Surveys were deleted randomly in order to obtain 20 surveys per group.

Of the 120 surveys used for Study II, 114 respondents provided information regarding highest degree obtained. Of these, 11% held master's degrees, 52% held master's degrees plus 30 additional graduate semester hours of credit, 28% held doctoral degrees, and 5% held other graduate degrees. The median year in which respondents received their terminal degree was 1980 (range = 1952 through 1993). Most of the respondents (80%) were employed as school psychologists, but 6% indicated that their primary role was an administrator or supervisor, 7% a university trainer, and 7% another role. Of the 120 surveys used, 23% were from northeastern, 27% from north-central, 26% from western, and 24% from southern parts of the United States. The sample characteristics for Study II are presented in Table 2.

Table 2

Sample Characteristics for Study II

Characteristic	n	%
Educational Level		
Master's	13	11
Master's + 30	47	41
Doctorate	48	42
Other	6	5
Region		
North East	28	23
North Central	32	27
West	31	26
South	29	24
Primary Role		
School Psychologist	91	80
Administrator/Supervisor	7	6
University Trainer	8	7
Other	8	7
Year of Terminal Degree		
1988-1993	26	24
1983-1987	17	15
1978-1982	21	19
1973-1977	25	23
1972 or before	21	19

A three (BR + AR; BR + IC; BR + AR + IC) by two (salient link; no link) multivariate analysis of variance (MANOVA) was conducted for the two dependent variables: (a) accuracy and (b) confidence. Results revealed a significant main effect for information, F (4, 226) = 6.45, $p < .001$, Wilks' Lambda statistic = .81. Neither the main effect for salient link, F (2, 113) = 1.32, $p > .05$, Hotellings generalized T-

Squared statistic = 2.67, nor the interaction between information and salient link, however, were significant, $F < 1$, Wilks' Lambda statistic = .99.

To investigate further the significant MANOVA main effect for information, separate univariate analyses of variance (ANOVA) were conducted for each of the dependent variables. These results showed that both accuracy and confidence contributed to the MANOVA results. For accuracy, results revealed a significant main effect for information, $F\ (2, 114) = 6.68$, $p < .01$. Neither the main effect for salience nor the interaction were significant ($p > .30$ in both cases). The power associated with salience for this analysis, however, proved to be inadequate to guard against type-II error (β = .85). An eta-squared statistic showed that information accounted for 11% of the variance in the accuracy rating. The two-way ANOVA summary table for accuracy is presented in Table 3.

Table 3

Two-way ANOVA Table for Accuracy

Source	SS	df	MS	F	p
Information	4989.72	2	2494.86	6.68	.002
Salience	340.03	1	340.03	.91	.34
Information by Salience	53.82	2	26.91	.07	.93
Error	42553.60	114	373.28		

Tukey post hoc comparisons for the significant main effect (i.e., information) revealed that the mean accuracy rating for the BR + IC condition (M = 25.87) was significantly lower (i.e., more accurate) than the BR + AR condition (M = 41.52), p < .01. The BR + AR + IC condition (M = 31.85) did not differ from either of the other two conditions, $p > .05$ in both cases. In other words, [BR + IC] < [BR + AR]; but [BR + IC] = [BR + AR + IC]; and [BR + AR] = [BR + AR + IC].

For confidence, ANOVA results revealed a significant main effect for information, F (2, 114) = 9.02, $p = .001$. Neither the main effect for salience nor the interaction were significant ($p > .10$ in both cases). Again, the power associated with salience for this analysis proved to be inadequate to guard against type-II error ($\beta = .71$). An eta-squared statistic showed that information accounted for 14% of the variance in the confidence rating. The two-way ANOVA summary table for confidence is presented in Table 4.

Table 4

Two-way ANOVA Table for Confidence

Source	SS	df	MS	F	p
Information	58.07	2	29.03	9.02	.001
Salience	7.50	1	7.50	2.33	.13
Information by Salience	1.40	2	.70	.22	.80
Error	367.00	114	3.22		

Tukey post hoc comparisons for the significant main effect (i.e., information) revealed that the mean confidence rating for the BR + IC condition ($M = 3.20$) was significantly lower (i.e., less confident) than either the BR + AR ($M = 4.70$) or the BR + AR + IC conditions ($M = 4.65$), $p < .01$ in both cases. The BR + AR condition did not differ from the BR + AR + IC condition, $p > .10$. In other words, [BR + IC] < [BR + AR] = [BR + AR + IC].

Analysis III

Next, because the main effect for salience and the interaction between salience and information were not significant in the two-way MANOVA, the two levels of salience (i.e., salient link; no link) were collapsed. A one-way MANOVA for four levels of information (i.e., base-rate information from Study I, BR; and BR + AR, BR + IC, and BR + AR + IC from Study II) was then conducted for the two dependent variables. Results revealed a significant effect for information, $F(6, 270) = 13.78$, $p < .001$, Wilks' Lambda statistic = .59.

To investigate further this significant MANOVA effect, ANOVAs again were conducted for each of the dependent variables. Similar to Analysis II, these results showed that both accuracy and confidence contributed to the MANOVA results. For accuracy, results revealed a significant effect for information, $F(3, 136) = 25.07$, $p < .001$. An eta-squared statistic showed that information accounted for 36% of the variance in the accuracy rating. The one-way ANOVA summary table for accuracy is presented in Table 5.

Table 5

One-way ANOVA Table for Accuracy

Source	SS	df	MS	F	p
Information	23752.69	3	7917.56	25.07	<.001
Error	42947.45	136	315.79		

Tukey post hoc comparisons for the significant effect (i.e., information) revealed that the mean accuracy rating for the base-rate condition (BR) from Study I ($M = 0$) was significantly lower (i.e., more accurate) than for any of the other three conditions, $p < .01$. Moreover, the mean accuracy rating for the BR + IC condition ($M = 26$) was significantly lower (i.e., more accurate) than the mean accuracy rating for the BR + AR condition (M = 42), p < .01. There was no difference, however, between the BR + AR + IC condition (M = 32) and the BR + IC or the BR + AR conditions, p > .05 in both cases. In other words, [BR] < [BR + IC] < [BR + AR]; but [BR + AR] = [BR + AR + IC]; and [BR + IC] = [BR + AR + IC]. Table 6 presents the mean probability ratings and the mean accuracy ratings (i.e., the absolute difference between correct and stated probability judgment) for each information condition.

Table 6

Mean Probability and Accuracy Ratings for Information Conditions

Information Condition	Probability Rating (M)	Accuracy Rating (M)
BR	10.00	.00
BR + IC	34.87	25.87
BR + AR	68.62	41.52
BR + AR + IC	48.70	31.85

For confidence, results also revealed a significant effect for information, $F(3, 136) = 10.05$, $p < .001$. An eta-squared statistic showed that information accounted for 18% of the variance in the confidence rating. The one-way ANOVA summary table for confidence is presented in Table 7.

Table 7

One-way ANOVA Table for Confidence

Source	SS	df	MS	F	p
Information	95.78	3	31.92	10.05	<.001
Error	432.10	136	3.17		

Tukey post hoc comparisons for the significant effect (i.e., information) revealed that the mean confidence ratings for both the BR ($\underline{M}$ = 2.70) and the BR + IC conditions ($\underline{M}$ = 3.20) were significantly lower (i.e., less confident) than either the BR + AR (M = 4.70) or the BR + AR + IC ($\underline{M}$ = 4.65) conditions, $\underline{p} < .01$. There were no significant differences between either the BR and BR + IC conditions or the BR + AR and BR + AR + IC conditions, $\underline{p} > .10$ in both cases. In other words, [BR] = [BR + IC] < [BR + AR] = [BR + AR + IC].

Analysis IV

Finally, an ANOVA for years since terminal degree was conducted across the four levels of case-scenario information in order to determine if differences in this dependent variable might account for any of the previous findings. Results, however, showed that year since terminal degree was distributed similarly across these four groups, $\underline{F}$ (3, 126) = 2.56, $\underline{p} > .05$.

Discussion

Based on research conducted with laypeople and professionals alike (e.g., Eddy, 1982; Kahneman & Tversky, 1973; Schwartz, Gory, Kassirer, & Essig, 1973), three major predictions were made. First, it was predicted that school psychologists would use base-rate information appropriately to make probabilistic diagnostic decisions in the absence of other clinical information. Second, it was predicted that providing a salient link between base-rate information and the diagnosis to be predicted would improve diagnostic accuracy. Finally, it was predicted that when school psychologists were presented with other clinical information (i.e., illusory-correlation and accuracy-rate information) in addition to base rates, a greater degree of error would be present when both individuating, that is, illusory-correlation (IC), and diagnostic, that is, accuracy-rate (AR) information were presented together with base-rates (BR), than when either of these sources of information was presented alone with base rates. That is, (BR + IC + AR) > (BR + IC) or (BR + AR).

Use of Base Rates

Results of the present investigation supported only the first hypothesis. Here, despite relatively low levels of diagnostic confidence, participants did use base-rate information appropriately in the absence of other clinical information. This clearly was shown in the responses of the school psychologists, every one of whom was completely accurate on the probabilistic diagnostic decision task that was administered (i.e., M difference between correct vs. stated probability = 0, SD = 0). This result is consistent with other research (Kahneman & Tversky, 1972; Lyon & Slovic, 1976) showing that when base rates are presented alone, individuals are able to use that information accurately. Thus, this result generalizes the tendency to use base-rate information appropriately in the absence of other clinical information to the sample of school psychologists used in the present study. It shows that the school psychologists appropriately used these relevant base-rate data when making probabilistic diagnostic decisions about a case. Although not surprising, this finding is important because it demonstrates that when no other clinical information is present, school psychologists (like most other professionals and lay persons) use base rates accurately.

Salience of Link between Base Rate and Diagnosis

The second hypothesis, stating that school psychologists would be more accurate in their probabilistic diagnostic decisions when a salient link was provided between the base-rate information and the category to be predicted, was not supported. Responses of school psychologists who were provided with a salient link in their case-scenario information did not differ significantly from those who did not receive a salient link. Thus, school psychologists were not more accurate in their probabilistic diagnostic decisions when they had a salient link provided between the base-rate information and the category to be predicted.

Other research examining salience as a potential intervention to remediate or correct the base-rate fallacy, has yielded equivocal results. Some research has documented that providing a salient link between the base-rate information and the category to be predicted enhances the relevance of the base rates and thereby improves accuracy (Bar-Hillel, 1980), whereas other attempts to improve accuracy by manipulating salience have proven to be unsuccessful (Lyon & Slovic, 1976). Because the power associated with this analysis was insufficient to guard adequately against type-II errors (i.e., $\beta = .85$), and because of the limited research conducted on salience especially within school psychology, salience currently should not be ruled out as a potential avenue for reducing the base-rate fallacy. Perhaps, a potentially more powerful link between the base-rate information and the category to be predicted may help to improve the perceived relevance of base-rate information and thereby increase diagnostic accuracy. This issue, of course, requires further empirical assessment.

Base-rate, Individuating, and Diagnostic Information

The third hypothesis, concerning differential levels of accuracy associated with combinations of base-rate, illusory-correlation (or individuating), and accuracy-rate (or diagnostic) information, was not supported. In contrast to what was predicted, psychologists who received both irrelevant illusory-correlation information as well as relevant accuracy-rate information with base rates were not the least accurate. Instead, their responses did not differ significantly from those who received either the relevant accuracy-rate information with base rates or those who received the irrelevant illusory-correlation information with base rates. Moreover, contrary to what might be expected, participants who received the irrelevant illusory-correlation information with base rates were more accurate than those who received the relevant accuracy-rate information with base rates.

This third hypothesis was based on previous research conducted in school psychology and cognitive psychology that has documented the effects of illusory-correlation and accuracy-rate information on probabilistic decision making. Specifically, research in the field of school psychology (Gnys et al., in press) has revealed that school psychologists' diagnostic decisions are influenced adversely by an illusory belief in an association between particular psychometric results and the diagnosis of a learning disability. Further, research in cognitive psychology (Lyon & Slovic, 1976) has documented that individuals often rely on the accuracy rate of a test and fail to incorporate other relevant information (e.g., base rates) when making probability judgments. Research has not examined how these two sources of information, when combined, affect diagnostic decision-making. It seemed reasonable, however, to speculate that diagnostic decision-making errors associated with both of these sources of information combined (i.e., both irrelevant illusory-correlation information and relevant accuracy-rate information) would be potentiated (or compounded) over those associated with either source of information alone. Thus, probabilistic diagnostic decisions for situations involving base-rate, illusory-correlation, and accuracy-rate information were predicted to be less accurate than for situations involving either (a) base-rate and illusory-correlation information or (b) base-rate and accuracy-rate information.

Possible explanations. As noted, however, this hypothesis was not supported in the present study. Moreover, the amount of error associated with adding relevant accuracy-rate information to base rates exceeded the amount of error associated with adding irrelevant illusory-correlation information to base rates. Results of a post-hoc analysis between information presented and year of terminal degree revealed that there were no significant differences according to this variable. Thus, year of terminal degree did not provide an explanation of the differential accuracy among the kinds of information presented. There are, however, several possible explanations for these differential levels of accuracy, including, but not limited to: (a) the extremity of the accuracy-rate information provided, (b) the substantative nature of the illusory-correlation information provided, and (c) the assumption of additive effects of error associated with these two sources of information.

The extremity of the accuracy-rate information (i.e., 80%) may provide one explanation for the differential levels of accuracy. For example, research has documented that when individuals are presented with accuracy-rate information, they often base their decisions on that information and neglect other pertinent information. Lyon and Slovic (1976) suggested that probability estimates often are determined by the most salient features of case-specific evidence. In this investigation, the stated accuracy-rate information may have been particularly salient given its extreme nature, that is, the procedure was described as one that is accurate in identifying the disorder as present 80% of the time, and inaccurate in identifying the disorder as present 20% of the time. Thus, participants may have made diagnostic decisions primarily on the basis of accuracy rate, thereby increasing error. Unfortunately, the design of this investigation did not permit an examination of the sources of information upon which participants relied in order to make their diagnostic decisions. Thus, an evaluation of this possibility awaits empirical verification.

The substantive nature of the illusory correlation may also provide an explanation of why school psychologists were more accurate when presented with base-rate and illusory-correlation information than with base-rate and accuracy-rate information. In the illusory-correlation conditions, participants were given the results of the WISC-R and were asked to make a probability judgment about the diagnosis of a learning disability. Contemporary definitions of learning disabilities rely on academic-achievement information (Hammill, 1990; Hooper & Willis, 1989), and the psychoeducational assessment of learning disabilities typically includes the administration of both aptitude and achievement tests. Because academic-achievement information was not included with aptitude information in these conditions, participants may have perceived that an essential source of diagnostic information was missing, thereby becoming more conservative (and, therefore, more accurate) in their probability estimates than participants in the base-rate and accuracy-rate conditions.

Finally, the assumption of additive effects of error associated with these two sources of information (i.e., accuracy-rate and illusory-correlation information) may have been incorrect. Instead, results of this study suggest that these two sources of information, when combined with base-rates, did not influence probability judgments in a solely additive fashion. Compounding or potentiating error would occur if these sources of information were independent, but error associated with one source of information also could offset error that is inversely correlated another source of information. Again, however, the design of this investigation did not permit an examination of this issue, because diagnostic-information conditions (e.g., BR + IC; BR + AR) were treated as a between-groups rather than as a within-subjects variable.

The base-rate fallacy. Because salience of the link between base-rate and diagnosis did not affect the accuracy of probabilistic diagnostic decision-making of the school psychologists, the two levels of this variable (i.e., salient link present vs. salient link absent) were collapsed, thereby permitting a four-way comparison among levels of diagnostic information. Here, in Analysis III, responses from Study I (involving only base-rate information) were compared with responses from Study II (involving BR + AR, BR + IC, and BR + AR + IC information). Results of this comparison revealed that participants who received base-rate information only were more accurate than all other participants. This finding is particularly important because it documented that when school psychologists were presented with base-rate information in combination with either irrelevant information, diagnostic information, or irrelevant and diagnostic information, as compared to base rates alone, their diagnostic accuracy decreased. This investigation, therefore, generalizes the base-rate fallacy to the present sample of school psychologists (Ajzen, 1977; Bar-Hillel, 1980; Lyon & Slovic, 1976; Nisbett & Borgida, 1975), and is perhaps the first study to demonstrate empirically that school psychologists are likely to neglect, or to underutilize, base rates in probabilistic decision tasks when either individuating information or diagnostic information is present.

Diagnostic Confidence

Results also revealed that participants presented with base-rate information only and those presented with base-rate and illusory-correlation information were less confident in their diagnostic judgments than those presented with base-rate and accuracy-rate information or with base-rate information combined with accuracy-rate and illusory-correlation information. There was no difference in confidence between participants who received base-rate information and those who received base-rate and illusory-correlation information. Similarly, there was no difference in confidence between participants who received base-rate and accuracy-rate information and those who received base-rate information combined with accuracy-rate and illusory-correlation information. That is, [BR] = [BR + IC] < [BR + AR] = [BR + AR + IC].

These results are consistent with other research that demonstrates that individuals who have considerable confidence in their decisions often make errors (Einhorn & Hogarth, 1978; Fischoff, 1977, 1982). Indeed, in this investigation, the most accurate school psychologists (i.e., base-rate only information) were less confident in their probabilistic diagnostic decisions than all groups but one (i.e., the base-rate plus illusory-correlation information group). Similarly, the participants who received illusory-correlation information reported lower confidence levels than the less accurate participants who received accuracy-rate information. These results suggest limited insight, in the sense that subjective levels of confidence did not match objective measures of accuracy for the present sample of school psychologists (cf., Fisch, Hammond, Joyce, & O'Reilly, 1981; Gauron & Dickinson, 1966).

Conclusions

Results of this investigation generalized the base-rate fallacy to the field of school psychology. School psychologists were able to use base-rate information accurately in the absence of other clinical information. Their accuracy, however, decreased when presented with additional information, either irrelevant or diagnostic. This finding is consistent with other research examining base rates, and highlights the need to develop interventions to address the neglect of base-rate information for diagnostic decision making.

The intervention designed to reduce the error in school psychologists probability judgments by increasing the salience of the base-rate information was not effective. The power of this analysis, however, was insufficient and therefore this variable requires addditional empirical evaluation.

Contrary to prediction, results also documented that combining irrelevant individuating and relevant diagnostic information did not generate the greatest amount of error in school psychologists' probability judgments. This result suggests that error, in fact, was not compounded, and raises questions about how different kinds of information are aggregated in probabilistic diagnostic decision-making tasks.

Finally, results documented that the more confident school psychologists were not the most accurate. Thus, the higher confidence of some of the participants in this study was unwarranted based on their accuracy. This result suggests limited insight about diagnostic decision processes.

Implications. It has been demonstrated that the neglect or underutilization of base rates can lead to serious diagnostic decision-making errors. Meehl and Rosen (1955) warned psychologists against evaluating tests by their accuracy rate alone; and, instead, emphasized the importance of a decision-making framework that incorporates base rates, costs of error, goals, and other pertinent information.

The findings of this investigation, combined with other decision-making research in school psychology (Aspel, 1992; Gnys et al., in press; Kennedy, Faust, Willis, & Piotrowski, in press) raise concerns about the diagnostic strategies of school psychologists. It appears that school psychologists are prone to some of the same decision-making errors (e.g., illusory correlation, lack of awareness, the base-rate fallacy) that have been documented in other fields (see Eddy, 1982; Schwartz et al., 1973).

Given that diagnosis is a probabilistic task, these results highlight the need for efforts to improve diagnostic accuracy by eliminating the base-rate fallacy. Clearly, school psychologists demonstrated that they use base rates when that information was presented in isolation. Unfortunately, in a more typical diagnostic decision-making task much more information usually is presented. Base rates typically are not presented alone and even may be obfuscated by poorly operationalized or poorly validated diagnostic criteria. Moreover, in many judgment tasks there is more non-predictive than predictive information. Sometimes information about the base rate of a diagnosis is unavailable to a clinician. In some cases, national base-rate information may be available, but information about base rates specific to the sample in question may be lacking. Despite its importance to diagnostic decision making, however, Faust and Nurcombe (1989) noted that little effort seems to have been directed towards collecting base-rate information.

Of course, individual clinicians can make base-rate estimates based on representative samples of cases, but this requires that they first be aware of the importance of base-rate information for diagnostic accuracy. Unfortunately, results of this study suggest that school psychologists have limited awareness of the importance of base-rate information for diagnostic accuracy. Other research that has attempted to improve accuracy by identifying problematic judgment strategies and warning diagnosticians to avoid them have proven unsuccessful (Arkes, 1981; Kurtz & Garfield, 1978), perhaps because awareness of a problematic decision-making strategy inevitably does not address the cognitive processes that underlie that strategy (Faust & Nurcombe, 1989).

Additional research is needed to explore potential avenues for improving the use of base-rate information in the presence of other clinical information. Results of this study present potential implications for trainers as well as practitioners. It has been suggested that one reason base rates may be neglected is because humans rely on heuristics to reduce the complexity of many cognitive tasks. Perhaps, providing alternative heuristics for probabilistic decision-making will enhance the use of base-rates. Faust and Nurcombe (1989) suggested several strategies to guide clinical decision making based in part on principles of probabilistic reasoning. Of course, much research is needed to see whether teaching these strategies will improve probabilistic decision making.

In conclusion, research in the area of probabilistic diagnostic making has revealed numerous problematic judgment strategies. There is now a growing body of research showing that the suboptimal reasoning strategies demonstrated in medicine, psychiatry, and clinical psychiatry also apply to the field of school psychology. This investigation generalized one of these judgment errors, that is, the base-rate fallacy, to diagnostic decisions made by school psychologists. Future research will need to begin to focus on corrective procedures and interventions to reduce diagnostic decision-making error.

Appendix A

Diagnostic Decision Making in School Psychology

We are asking the participation of professionals involved in school psychology to help us in an investigation of diagnostic decisions. Our aim is to improve our understanding of school psychologists' diagnostic decisions. We believe your participation can make an important contribution to research in school psychology.

We would greatly appreciate you taking a few minutes to review the case on the next page and to respond to a few items. Data will be analyzed on a group basis. We do guarantee anonymity and wish to assure you that the code number that appears on the page will be used only to avoid duplicate mailings. A return envelope is enclosed for your convenience. If you are interested, we would be happy to send you a mailing indicating when and where any publications stemming from this work will appear. Just send us a note indicating your interest under separate cover (to ensure anonymity).

Thanks for considering this request.

Sincerely,

W. Grant Willis

W. Grant Willis, PhD,
Associate Professor
University of Rhode Island

Mary Lynne Kennedy

Mary Lynne Kennedy, MA
Graduate Student
University of Rhode Island

PART I. Professional/Practice Characteristics:

1. Highest Degree (check one):

_____BA _____EdS

_____MA, MS, MEd _____EdD, PhD, PsyD

_____Masters +30 _____other (specify):___________

2. Year of degree:______

3. Gender (circle): Male Female

4. State in which you are employed:________________

5. Primary Role:

_____School psychologist _____University Trainer

Administrator/Supervisor Other (specify):

Appendix B

Part II. Diagnostic Decision Making:

Please read the case described below and answer the proceeding questions.

(1) A total of 10% of the boys in John Smith's school have a learning disability.

1.) Please indicate the probability that John Smith has a learning disability on a scale from 0% (lowest -- not probable at all) to 100% (highest -- certain):

There is a _____% that this child is learning disabled.

2.) Please rate your confidence in your decision on a scale of 1 (not at all confident) to 7 (very confident):

Not At All Confident ___ ___ ___ ___ ___ ___ ___ Very Confident
1 2 3 4 5 6 7

Part II. Diagnostic Decision Making:

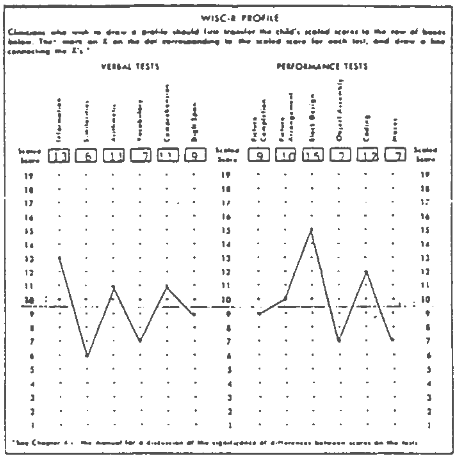
Please read the case described below and answer the proceeding questions.

(1) A procedure is used to diagnose learning disabilities.

(2) Because of an effective early intervention program in John Smith's school, the incidence of learning disabilities among boys in this school is only 10%.

(3) John Smith has been tested and results are positive for a learning disability.

(4) Part of the procedure for diagnosing learning disabilities involves administering an intelligence test. Below you will find John Smith's WISC-R (Wechsler Intelligence Scale for Children--Revised) results.



	Scaled Score
VERBAL TESTS	
Information	13
Similarities	6
Arithmetic	11
Vocabulary	7
Comprehension	11
(Digit Span)	9
Verbal Score	48
PERFORMANCE TESTS	
Picture Completion	9
Picture Arrangement	10
Block Design	15
Object Assembly	7
Coding	12
(Mazes)	7
Performance Score	53

	Scaled Score	IQ
Verbal Score	48	97
Performance Score	53	104
Full Scale Score	101	100

*Prorated from 4 tests, if necessary

1.) Please indicate the probability that John Smith has a learning disability on a scale from 0% (lowest -- not probable at all) to 100% (highest -- certain):

There is a _____% that this child is learning disabled.

2.) Please rate your confidence in your decision on a scale of 1 (not at all confident) to 7 (very confident):

Part II. Diagnostic Decision Making:

Please read the case described below and answer the proceeding questions.

(1) A procedure is used to diagnose learning disabilities.

(2) When the results of the procedure are positive for a learning disability, the procedure is correct 80% of the time in identifying learning disabled students as learning disabled. The procedure is wrong 20% of the time, that is, it misidentifies children who are not learning disabled as learning disabled.

(3) A total of 10% of the boys in John Smith's school have a learning disability.

(4) John Smith has been tested and results are positive for a learning disability.

1.) Please indicate the probability that John Smith has a learning disability on a scale from 0% (lowest -- not probable at all) to 100% (highest -- certain):

There is a _____% that this child is learning disabled.

2.) Please rate your confidence in your decision on a scale of 1 (not at all confident) to 7 (very confident):

Not At All Confident ___ ___ ___ ___ ___ ___ ___ Very Confident

1 2 3 4 5 6 7

Part II. Diagnostic Decision Making:

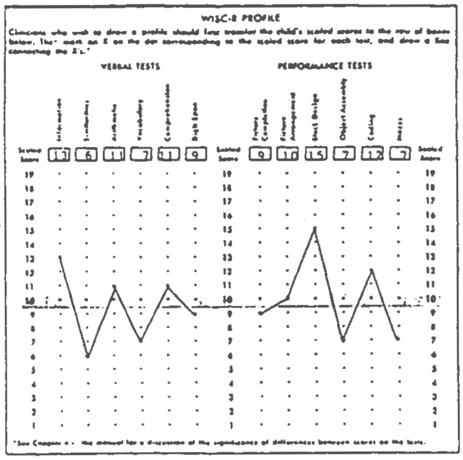
Please read the case described below and answer the proceeding questions.

(1) A procedure is used to diagnose learning disabilities.

(2) A total of 10% of the boys in John Smith's school have a learning disability.

(3) John Smith has been tested and results are positive for a learning disability.

(4) Part of the procedure for diagnosing learning disabilities involves administering an intelligence test. Below you will find John Smith's WISC-R (Wechsler Intelligence Scales for Children--Revised) results.



	Scaled Score
VERBAL TESTS	
Information	13
Similarities	6
Arithmetic	11
Vocabulary	7
Comprehension	11
(Digit Span)	(9)
Verbal Score	48
PERFORMANCE TESTS	
Picture Completion	9
Picture Arrangement	10
Block Design	15
Object Assembly	7
Coding	12
(Mazes)	(7)
Performance Score	53

	Scaled Score	IQ
Verbal Score	48	*97
Performance Score	53	*104
Full Scale Score	101	100

*Prorated from 4 tests, if necessary.

1.) Please indicate the probability that John Smith has a learning disability on a scale from 0% (lowest -- not probable at all) to 100% (highest -- certain):

There is a _____% that this child is learning disabled.

2.) Please rate your confidence in your decision on a scale of 1 (not at all confident) to 7 (very confident):

Part II. Diagnostic Decision Making:

Please read the case described below and answer the proceeding questions.

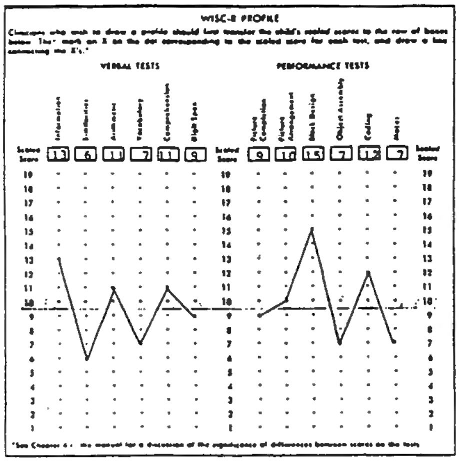
(1) A procedure is used to diagnose learning disabilities.

(2) When the results of the procedure are positive for a learning disability, the procedure is correct 80% of the time in identifying learning disabled students as learning disabled. The procedure is wrong 20% of the time, that is, it misidentifies children who are not learning disabled as learning disabled.

(3) A total of 10% of the boys in John Smith's school have a learning disability.

(4) John Smith has been tested and results are positive for a learning disability.

(5) Part of the procedure for diagnosing learning disabilities involves administering an intelligence test. Below you will find John Smith's WISC-R (Wechsler Intelligence Scales for Children--Revised) results.



	Scaled Score
VERBAL TESTS	
Information	13
Similarities	6
Arithmetic	11
Vocabulary	7
Comprehension	11
(Digit Span)	(9)
Verbal Score	48
PERFORMANCE TESTS	
Picture Completion	9
Picture Arrangement	10
Block Design	15
Object Assembly	7
Coding	12
(Mazes)	(7)
Performance Score	53

	Scaled Score	IQ
Verbal Score	48	* 97
Performance Score	53	* 104
Full Scale Score	101	100

*Prorated from 4 tests, if necessary.

1.) Please indicate the probability that John Smith has a learning disability on a scale from 0% (lowest -- not probable at all) to 100% (highest -- certain):

There is a _____% that this child is learning disabled.

2.) Please rate your confidence in your decision on a scale of 1 (not at

Part II. Diagnostic Decision Making:

Please read the case described below and answer the proceeding questions.

(1) A procedure is used to diagnose learning disabilities.

(2) When the results of the procedure are positive for a learning disability, the procedure is correct 80% of the time in identifying learning disabled students as learning disabled. The procedure is wrong 20% of the time, that is, it misidentifies children who are not learning disabled as learning disabled.

(3) Because of an effective early intervention program in John Smith's school, the incidence of learning disabilities among boys in this school is only 10%.

(5) John has been tested and results are positive for a learning disability.

1.) Please indicate the probability that John Smith has a learning disability on a scale from 0% (lowest -- not probable at all) to 100% (highest -- certain):

There is a _____% that this child is learning disabled.

2.) Please rate your confidence in your decision on a scale of 1 (not at all confident) to 7 (very confident):

Not At All Confident ___ ___ ___ ___ ___ ___ ___ Very Confident

1 2 3 4 5 6 7

Part II. Diagnostic Decision Making:

Please read the case described below and answer the proceeding questions.

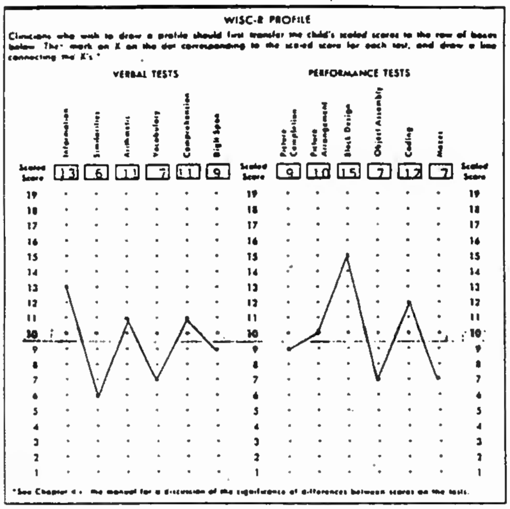
(1) A procedure is used to diagnose learning disabilities.

(2) When the results of the procedure are positive for a learning disability, the procedure is correct 80% of the time in identifying learning disabled students as learning disabled. The procedure is wrong 20% of the time, that is, it misidentifies children who are not learning disabled as learning disabled.

(3) Because of an effective early intervention program in John Smith's school, the incidence of learning disabilities among boys in this school is only 10%.

(4) John Smith has been tested and results are positive for a learning disability.

(5) Part of the procedure for diagnosing learning disabilities involves administering an intelligence test. Below you will find John Smith's WISC-R (Wechsler Intelligence Scales for Children--Revised) results.



	Scaled Score
VERBAL TESTS	
Information	13
Similarities	6
Arithmetic	11
Vocabulary	7
Comprehension	11
(Digit Span)	(9)
Verbal Score	48
PERFORMANCE TESTS	
Picture Completion	9
Picture Arrangement	10
Block Design	15
Object Assembly	7
Coding	12
(Mazes)	(7)
Performance Score	53

	Scaled Score	IQ
Verbal Score	48	* 97
Performance Score	53	* 104
Full Scale Score	101	100

*Prorated from 4 tests, if necessary

1.) Please indicate the probability that John Smith has a learning disability on a scale from 0% (lowest -- not probable at all) to 100% (highest -- certain):

There is a _____% that this child is learning disabled.

2.) Please rate your confidence in your decision on a scale of 1 (not at all confident) to 7 (very confident):

Bibliography

Ajzen, I. (1977). Intuitive theories of events and the effects of base rate information on prediction. Journal of Personality and Social Psychology, 35, 303-314.

Arkes, H. R. (1989). Principles in judgment/decision-making: Research pertinent to legal proceedings. Behavioral Sciences & the Law, 7, 429-456.

Arkes, H. R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. Journal Of Consulting and Clinical Psychology, 49, 323-330.

Arkes, H. R., & Harkness, A. R. (1980). Effect of making a diagnosis on subsequent recognition of symptoms. Journal of Experimental Psychology, 6, 568-575.

Aspel, A. (1992). School psychologists' awareness of their decision-making processes. Unpublished master's thesis, University of Rhode Island, Kingston.

Bar-Hillel, M. (1975). The base rate fallacy in probability judgments. Unpublished doctoral dissertation, Hebrew University, Jerusalem.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. Acta Psychologica, 44, 211-233.

Bar-Hillel, M. and Fischoff, B. (1981). When do base rates affect predictions? Journal of Personality and Social Psychology, 41, 671-680.

Borgida, E., & Nisbett, R. E. (1977). The differential impact of abstract vs. concrete information on decisions. Journal of Applied Social Psychology, 7, 258-271.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. Journal of Abnormal Psychology, 72, 193-204.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases (pp. 249-267). New York: Cambridge University Press.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. Psychological Review, 85, 86-106.

Faust, D., & Nurcombe, B. (1989). Improving the accuracy of clinical judgment. Psychiatry, 52,197-208.

Fisch, H., Hammond, K., Joyce, C., & O'Reilly, P. (1981). An experimental study of the clinical judgment of general physicians in evaluating and prescribing for depression. British Journal of Psychiatry, 138, 100-109.

Fischoff, B. (1977). Perceived informativeness of facts. Journal of Experimental Psychology: Human Perception and Performance, 3, 349-358.

Fischoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic & A. Tversky, (Eds.), Judgment under uncertainty: Heuristics and biases (pp. 422-444). New York: Cambridge University Press.

Gauron, E. F., & Dickinson, J. K. (1966). Diagnostic decision-making in psychiatry. Archives of General Psychiatry, 14, 225-237.

Gnys, J., Willis, W. G., & Faust, D. (in press). School psychologists' diagnoses of learning disabilities: A study of illusory correlation. Journal of School Psychology.

Hammill, D. D. (1990). On defining learning disabilities: An emerging consensus. Journal of Learning Disabilities, 23, 74-84.

Hamilton, D. L., & Rose, T. (1980). Illusory correlation and the maintenance of stereotypic beliefs. Journal of Personality and Social Psychology, 39, 832-845.

Hooper, S. R., & Willis, W. G. (1989). Learning disability subtyping: Neuropsychological foundations, conceptual models, and issues in clinical differentiation. New York: Springer-Verlag.

Israel, J., & Tajfel, H. (Eds.). (1972). The context of social psychology: A critical assessment. London: Academic Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology, 3, 430-454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. Psychological Review, 80, 237-251.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. Cognition, 11, 123-141.

Kaufman, A. S. (1976). A new approach to the interpretation of test scatter on the WISC-R. Journal of Learning Disabilities, 9, 160-168.

Kaufman, A. S. (1979). Intelligent testing with the WISC-R. New York: Wiley-Interscience.

Kennedy, M. L., Faust, D., Willis, W. G., & Piotrowski, C. (in press). Social-emotional assessment practices of school psychologists: A national survey. Journal of Psychoeducational Assessment.

Kiriwan, J. R., Chaput de Saintonge, D. M., Joyce, C. R. B., & Currey, L. F. (1983). Clinical judgment in rheumatoid arthritis II: Judging "current disease activity" in clinical practice. Annals of Rheumatic Diseases, 42, 648-651.

Kurtz, R. M., & Garfield, S. L. (1978). Illusory correlation: A further exploration of Chapman's paradigm. Journal of Consulting and Clinical Psychology, 46, 1009-1015

Lyon, D., & Slovic, P. (1976). Dominance of accuracy information and neglect of base rates in probability estimation. Acta Psychologica, 40, 287-298.

McDermott, P. A., Gluttin, J. J., Jones, J. N., Watkins, M. W., & Kush, J. (1989). Core profile types in the WISC-R national sample: Structure, membership, and applications. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 4, 292-299.

Meehl, P. E. (1954). Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota Press.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. Psychological Bulletin, 52, 194-215.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

National Association of School Psychologists. (1991). Directory of nationally certified school psychologists. Silver Springs, MD: Author.

Neisser, U. (1976). Cognition and reality: Principles and implications of cognitive psychology. San Francisco: Freeman.

Nisbett, R. E., & Borgida, E. (1975). Attribution and the psychology of prediction. Journal of Personality and Social Psychology, 32, 932-943.

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

Reynolds, C. R., & Kaufman, A. S. (1990). Assessment of children's intelligence with the Wechsler Intelligence Scale for Children-Revised (WISC-R). In C. R. Reynolds & R. W. Kamphaus (Eds.), Handbook of psychological and educational assessment of children: IQ and achievement (pp. 127-165). New York: Guilford Press.

Schneider, D. (1988). Social psychology. New York: Harcourt Brace Jovanovich, Inc.

Schwartz, W. B., Gory, G. A., Kassirer, J. P., & Essig, A. (1973). Decision analysis and clinical judgment. American Journal of Medicine, 55, 459-472.

Smedslund, J. (1963). The concept of correlation in adults. Scandinavian Journal of Psychology, 4, 165-173.

Tajfel, H. (1969). The cognitive aspects of prejudice. Journal of Social Issues, 25, 79-94.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Tversky, A., & Kahneman, D. (1980). Causal thinking in judgment under uncertainty. In M. Fishbein (Ed.), Progress in social psychology. Hillsdale, NJ: Erlbaum.

Ward, W. D., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. Canadian Journal of Psychology, 19, 231-241.

Ziskin, J. & Faust, D. (1991). Coping with psychiatric and psychological testimony (Vol. 1). Los Angeles, CA: Law and Psychology Press.