HOLISTIC JUDGMENT IN CLINICAL PRACTICE
Utility or Futility?

John Ruscio
Elizabethtown College

Clinical decision making sometimes proceeds not by independently considering each of a number of relevant variables, but by evaluating a complex whole. Such a holistic approach, premised on the notion that "everything influences everything else," raises logical and psychological difficulties. Reaching truly holistic judgments requires knowledge of far more unique configurations of data than have ever existed, demands feats of information integration that are more incompatible with our understanding of human cognitive limitations, and is ordinarily unwarranted in light of the relationships between most variables and phenomena of interest. In its strong form, holism constitutes an approach to professional practice that is unnecessarily and unattainably complex, often misguided, and potentially unethical by virtue of the existence of demonstrably superior approaches to decision making. Speculations are offered for conditions under which limited forms of holism may be useful.

Clinical assessment, case formulation, diagnosis, treatment planning, and therapy yield an astounding amount of client information, and it can be exceedingly difficult to integrate the relevant data to reach judgments and decisions of the greatest clinical utility. Holistic judgments are premised on the notion that interactions among all of the information must be taken into account to properly contextualize data gathered in a realm where everything can influence everything else. In contrast to strategies that more closely resemble an additive sum of main effects, some of the most popular approaches to personality assessment espouse holistic judgment. For example, whereas the validity of inferences drawn from individual clinical scales of the Minnesota Multiphasic Personality Inventory rests on a rigorous approach to scale development through empirical criterion keying (Hathaway & McKinley, 1940), there is comparatively meager support for the more holistic profile interpretations based on two- or three-point "code types" (see Graham, 2000). Likewise, individuals who use an assessment tool with dubious norms and psychometric properties—such as the Rorschach ink blot test (Lilienfeld, Wood, & Garb, 2000; Wood, Lilienfeld, Nezworski, & Garb, 2001)—sometimes argue that their interpretations are nonetheless valid because they are supplemented by consideration of additional information (e.g., Merlo & Barnett, 2001) that is integrated in a holistic fashion. Such appeals to holism presume that clinicians are adept at filtering and interactively combining massive amounts of data.

Two central questions, then, involve just how well human judgment can handle a plethora of information that involves high-order interaction effects and what the utility of such a reasoning process would be in clinical practice. The present paper addresses these questions through a uniquely multi-faceted evaluation of holistic judgment that draws upon its logical and psychological implications and reviews evidence from several relevant lines of research. First, it is argued that truly holistic judgment often entails logical and psychological impossibilities. Second, the performance of human judgment with complex tasks is shown to provide a poor empirical foundation for claims that clinicians can integrate information in a holistic manner. Third, the utility of holistic judgment is challenged on the grounds that monotonic interaction effects are captured remarkably well even by simple lin-

John Ruscio, Department of Psychology, Elizabethtown College.

Correspondence concerning this article should be addressed to John Ruscio, Department of Psychology, Elizabethtown College, One Alpha Drive, Elizabethtown, PA 17022. E-mail: rusciojp@etown.edu.

ear models containing only additive main effects. Fourth and finally, the questionable ethical foundation of practicing or advocating holistic judgment is explored in light of alternative approaches to decision making that achieve demonstrably superior clinical utility.

WHAT IS HOLISTIC JUDGMENT?

A half century ago, psychologists were embroiled in a debate over the utility of unaided human judgment for combining information to reach important decisions. In what remains the definitive treatise on this subject, Meehl (1954) argued that there is no theoretical or empirical reason to suspect that we can combine information in our heads as effectively as we can by using a simple statistical prediction rule (SPR; Swets, Dawes, & Monahan, 2000) or actuarial procedure. Over the past 50 years, a large body of research has accumulated to show that Meehl was correct (e.g., Dawes, Faust, & Meehl, 1989; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Grove & Meehl, 1996). Despite these data, some clinicians still prefer to use their heads rather than an SPR (Meehl, 1993). One of the reasons underlying professionals' adherence to clinical prediction methods may be the belief that mechanical procedures such as SPRs and actuarial tables cannot take into account all relevant material in the complex manner often thought to be necessary in applied clinical work. Practitioners may believe that their judgment processes accommodate a wider range of relevant information and integrate it in more sophisticated ways. Holism may represent one manifestation of this larger belief that is widespread in the mental health professions (Meehl, 1986).

Early studies in the clinical-statistical prediction literature led to speculations that whereas human judgment fares poorly against SPRs for making simple decisions, it might be superior when the information under consideration is more complex. Clinicians were expected to prevail with task characteristics involving configural, interactive relationships among variables (Meehl, 1954, 1967). Such holistic judgment cannot be reduced to—or reproduced by—additive main effects, for such effects would fail to adequately contextualize the available information. Rather, to reason holistically one must consider each piece of information in light of all available information, which requires processing based on interaction effects. If the meaning of each datum truly depends upon all of the other available information, then judgment would seem to require an interaction term of the highest order, not an additive function of the individual

factors. Thus, the present evaluation will deal with holistic judgment in its strong form: a consideration of high-order interaction effects among all available information. I will briefly consider more limited forms of holistic judgment toward the conclusion.

LOGICAL AND PSYCHOLOGICAL IMPOSSIBILITIES

Like other adherents to the clinical approach to prediction, holists may be overestimating the extent to which people can validly combine information in their heads. Assertions that they can take into account a vast array of information in a complex manner ignore the cognitive limitations of all humans. In their appeal to holism, practitioners claim to be able to work with an astounding number of unique configurations of information, often far exceeding even the number of people who have ever lived. This poses a serious problem for the logic of prediction as well as the psychological possibility of possessing or accessing a sufficiently comprehensive and finely differentiated knowledge base.

For example, consider how this difficulty manifests itself in modern astrology, which involves a problematic presumption regarding the number of unique horoscopes that exist. Astrology is based on the belief that all natural phenomena are influenced by celestial bodies. There is partial truth to this belief: "Every time we wake up with the sun, or plan barbecues on moonlit nights, or go fishing at high tide, we are showing how celestial bodies have real influence in our lives" (Kelly, 1997, p. 1035). But these influences are far more trivial than the grand claims of astrology suggest. For centuries, astrologers were consulted for assistance in important, practical decisions: Should I marry this person? Should I wage war against this nation? Around the 1950s, however, science began to catch up with astrology and test the predictions that astrologers made from factors in the horoscope. These tests revealed that astrologers' predictions were no more valid than chance-level guessing (for reviews, see Dean, Mather, & Kelly 1996; Kelly, 1997, 1998).

In the face of overwhelming negative evidence, many astrologers began to embrace holism by asserting that they must make use of the *whole chart* to render professional judgments. Thus, modern astrology often pays little attention to the evidence against it because "the horoscope is a whole system in which every part is influenced by every other part" (Kelly, 1997, p. 1044). Scientific investigations are often ignored by astrologers because they have tested only one astrological factor at a time, thereby failing to take into account the whole chart

that professional astrologers supposedly utilize to make their predictions. However, this holism defense is fatally flawed. One can rightly ask just what the whole chart means, and how many unique charts exist:

> Where does the whole chart end? With ten planets [astrologers count the sun and our moon, though none of the dozens of other moons in our solar system, as planets], twelve signs, twelve houses, midpoints, Arabic points, nodes, aspects and whatever other astrological concepts may be used, it is simply impossible to interpret a "whole chart." When astrologers claim that they use the whole chart, they only refer to the fact that they use more factors than just one. Nevertheless, no matter how many factors they use, they always use a restricted number of factors, and therefore only a part of the horoscope. They never use the whole chart. But then the question becomes how many factors would be considered, and which factors? ... Suppose that I consider as many as 20 factors, then undoubtedly an astrologer will come up who claims that I should use 21 factors. (Van Rooij, 1994, p. 56)

Based on a typical list of astrological factors, Kelly (1997) calculated that there are approximately 10^{28} possible horoscopes. Without providing details for her calculations, one astrologer arrived at the figure of 5.39×10^{68} (Doane, 1956). Either of these values undermines the foundation of holism, because there are far more unique horoscopes than the number of people who have ever lived. Thus, it is entirely possible—and, depending on which number of horoscopes one accepts, perhaps even highly probable—that most every person ever born has had a unique horoscope. This means that should you consult an astrologer, she would have neither met nor learned anything about another person with your horoscope. If nobody has ever had your unique horoscope before, how could an astrologer know what to predict for your future? The individual factors in your chart are the only usable information—they *are* shared with other people, and can therefore establish a basis for prediction—but holistic astrologers cannot consider one factor at a time because they would not be using the whole chart. There is no logical basis on which to make predictions in the complete absence of prior knowledge of similar cases. With a sufficiently rich array of information, the "everything influences everything else" nature of holism invariably leads to this crippling paradox.

In clinical work, the amount of available information is surely no less numerous and complex than the factors in an astrological horoscope. Moreover, it may be even more problematic for holistic judgment for several reasons. The sheer volume of individuating factors provided by a single client is staggering. In assessment, case formulation, diagnosis, treatment planning, and other important tasks that are inherently predictive in nature, a clinician may have to consider presenting complaints and referral information; signs and symptoms of one or more psychological disorders; scores on standardized or unstandardized psychological tests; information gleaned from structured, semistructured, or unstructured interviews; family, medical, psychiatric, and other history data; client and informant reports of social and occupational stressors; informal observations of client mannerisms and behavior; and so forth. If indeed the meaning of each piece of information depends upon all other available information, there may well be at least as many unique clinical presentations possible as there are horoscopes.

To make matters even more challenging, the clinician must consider not only factors that *are* disclosed or otherwise made available, but also factors whose *absence* is informative. Relevant data may be absent for a number of reasons, including a failure of the clinician to inquire, inaccessible or inaccurate client memories, and incomplete or deceptive reporting by the client. Whereas astrologers could, in principle, reach a consensus of opinion on a set of factors necessary and sufficient for valid interpretations and methods for assessing these factors, clinicians may never be so readily assured that they have all necessary information. Thus, clinicians who appeal to the virtues of holistic judgment place themselves in at least as untenable a position as do modern astrologers. The knowledge base required for truly holistic judgments—those based purely on interacting, rather than additive factors—is most unlikely to exist. Even if it did, it would be psychologically impossible to store and access such an incomprehensibly vast knowledge base.

RESEARCH ON THE PREDICTIVE VALIDITY OF CLINICAL JUDGMENT

Research in the tradition of comparing clinical and statistical prediction has direct bearing on claims regarding holistic judgment and suggests more practical alternatives. The accuracy of judgments made in a methodical way from just a few relevant pieces of information is usually equal or superior to those of experts who combine a wide array of information using unaided human judgment (for reviews and a meta-analysis, see Dawes et al., 1989; Grove et al., 2000; Grove & Meehl, 1996). Indeed, to date there is no replicable counterexample to this empirical generalization.

Explaining the Superiority of SPRs

This superiority of SPRs over clinical judgment has been attributed to two complementary sources: the desirable mathematical properties of SPRs and the cognitive limitations and biases of human judgment. SPRs weight information according to its empirical validity, predict with perfect reliability, work well regardless of the units of measurement of individual predictors, and take both the redundancy between predictors and statistical regression into account (Goldberg, 1991). In contrast, because our memory and processing capabilities are limited, our judgment often relies on mental shortcuts or heuristics (Kahneman, Slovic, & Tversky, 1982) such as availability (the easier we can recall or imagine instances of an event, the more common we judge it to be; Ruscio, 2000a; Tversky & Kahneman, 1973) or representativeness ("like goes with like"; Gilovich & Savitsky, 1996; Kahneman & Tversky, 1972). Heuristics such as these can lead to misjudged probabilities or frequencies and the creation of superstitious beliefs or illusory correlations (Chapman & Chapman, 1967, 1971; Kurtz & Garfield, 1978). We are also strongly biased in favor of our prior beliefs and are adept at constructing post hoc explanations (Ruscio, 1998b) that sacrifice historical truth for narrative truth. All of these biases, as well as many others (e.g., Kahneman, Slovic, & Tversky, 1982; Nisbett & Ross, 1980; Turk & Salovey, 1988), can contribute to a poor use of information, especially relative to SPRs.

Available evidence suggests that unaided human judgment cannot compete with a more mechanical process that involves a comparatively simple combination of a small handful of relevant variables (Swets et al., 2000). This conclusion has been supported in a tremendous number of disciplines, including a wide range of decisions made by trained and experienced professionals (Dawes, Faust, & Meehl, 1993): differential diagnoses of medical conditions; predictions of the longevity of chronically ill patients; predictions of success that lead to the acceptance or rejection of applicants to colleges, graduate programs, or jobs; predictions of dangerousness in parole hearings; predictions of the outcomes of sporting events used to set gambling odds; and predictions underlying financial transactions such as lending money and issuing insurance policies.

When provided with identical information, SPRs tend to achieve greater empirical accuracy than do professionals. This remains true when one provides professionals with information not available to the SPR, and even when one provides the results of the SPR itself, in which case professionals identify too many "exceptions"

to the rule (Dawes et al., 1989). Following Meehl's (1954) discussion of the latter problem, such overidentified counterexamples are often referred to as "broken leg" cases. For half a century, there has been a concerted, though so far unsuccessful, research effort to find even a single domain in which clinical judgment consistently surpasses the accuracy of statistical decision making.

The evidence is also robust in another sense: several different types of linear equations can outpredict human judges. That is, not only are SPRs that optimally weight information equal or superior to clinical judgment, but so are those that preserve only the direction of relationships—positive or negative—and weight the predictors equally (Dawes, 1979). Moreover, research shows that judgments are more accurate when made from a limited number of valid predictors; extra information that *could* be ignored typically is not, which serves to dilute the quality of judgments (Nisbett, Zukier, & Lemley, 1981; Ruscio, 2000b). The demonstrated superiority of unit-weighted models over judges prompted Dawes and Corrigan (1974) to conclude that "the whole trick is to know what variables to look at and then to know how to add" (p. 105). Unfortunately, holists often refuse to narrow their attention to a common set of demonstrably relevant variables, believing as they do that they must adequately contextualize each and every bit of data. They also insist that a far more complex process than mere addition is necessary to make accurate judgments and reach wise decisions, despite considerable evidence to the contrary.

The Search for Configural Judges

Just as the predictive failure of individual factors within the horoscope forced astrologers to retreat to holism, early research on the clinical-statistical prediction controversy led many people to speculate that clinical judgment would prevail at a more complex task. Unaided human judgment was acknowledged to be inferior to a SPR when making simple decisions, but was hypothesized to be superior when the information itself was interrelated in complex ways. As in the context of holism, it was proposed that task characteristics involving such configural relationships as interactions among variables might favor the clinical practitioner (Meehl, 1954, 1967).

Although there is reliable evidence that clinicians engage in *nonlinear* processing (e.g., Ganzach, 1994, 1995, 2001)—particularly with cognitively straightforward strategies that involve conjunctive or disjunctive rules (Dawes, 1964; Einhorn, 1971)—clinicians do not appear to draw on *interactions* to any substantial degree (Goldberg, 1968, 1991; Slovic & Lichtenstein, 1971;

Stewart, 1988). This result makes intuitive sense, for it is odd to presume that although clinical judgment is demonstrably poor at the relatively simple tasks that have been studied (e.g., making decisions based on a small number of valid predictors that are all linearly related to the outcome), it will function superbly with *more complex* tasks (e.g., making decisions based on a large number of variables, each of different and questionable validity, that interact with one another to an unknown extent). If a mechanical process can eke out an advantage even with simple tasks, it will likely far outpace human judgment with more complex tasks. In fact, research suggests that we are able to work effectively with up to about eight pieces of information at once (Cooksey, 1996), and there is absolutely no evidence that we are capable of psychologically manipulating even this much information if it is interactive.

In two recent experiments (Ruscio & Stern, 2001), individuals' ability to make holistic judgments was evaluated directly by explicitly showing them how to do so and then testing their performance. Participants were provided with full specifications for a relatively simple judgment task: data on two quantitative variables were available for each of a series of cases, and participants were asked to predict a quantitative criterion. All variables were normally distributed as T scores, and participants were told and shown graphically what this means. Participants were given specific instructions on how to generate accurate predictions: in one experimental condition, two factors were *additively* related to the criterion, whereas in another experimental condition, two factors were *interactively* related to the criterion (for simplicity, the pattern was a pure crossover interaction). The extent to which the criterion was predictable from the two factors was held constant across conditions, and each participant was instructed to adopt a judgment process in accordance with his experimental condition (those in the additive condition were told to use an additive process, whereas those in the interactive condition were told to use an interactive process) to maximize accuracy. Consistency was assessed through the reproducibility of judgments from the sum of two main effects (in the additive condition) or the sum of two main effects plus an interaction term (in the interactive condition), and accuracy was evaluated via the correlation between each participants' judgments and the criterion scores for his experimental condition.

Relative to judgments made in the additive condition, judgments made in the interactive condition were inconsistent (mean Rs of .65 vs. .79) and highly inaccurate (mean rs of .18 vs. .50). A second experiment repli-

cated these results and extended them across participants' education levels (freshmen and sophomores, juniors and seniors, faculty) and academic disciplines (social sciences, physical and natural sciences, professional studies, arts and humanities). The inability of these individuals to make consistent and accurate holistic judgments in a task of minimal complexity underscores the cognitive limitations that severely constrain the psychological feasibility of truly holistic judgment based on interactive effects.

THE MISGUIDED APPLICATION OF HOLISTIC JUDGMENT IN CLINICAL PRACTICE

In addition to the psychological difficulties already described, holistic judgment may often be unnecessary or helpful in clinical practice because an additive sum of main effects can be astonishingly potent in predicting outcomes. Dawes (1979) argued that naturally occurring relationships between psychological variables and outcomes of interest tend to be *monotonic*. That is, the direction of a variable's effect does not typically change as it interacts with other variables. For example, if a form of social skills training is more effective for extraverts than for introverts, and this gap widens with increasing impairment in social functioning, there is a monotonic personality × impairment interaction in predicting outcome (see Figure, top panel). On the other hand, if the training is more effective for extraverts than introverts at high levels of impairment, but more effective for introverts than for extraverts at low levels of impairment, there is a crossover interaction (see Figure, middle panel). Whereas the crossover interaction between personality and degree of impairment defies approximation through additive main effects, the monotonic interaction can be approximated quite well. The dotted lines in the bottom panel of the Figure show the predictions generated by additive main effects. Indeed, an interaction effect would provide little incremental validity. This is generally true of monotonic interactions, for the interactive component of the relationship between predictors and criterion is small relative to the main effect components.

Consider a hypothetical numerical example with just three interacting variables (i, j, and k; Yntema & Torgerson, 1961). If each factor varies as a rectangular distribution along the integers from 1 to 7 (yielding a data set of $7 \times 7 \times 7 = 343$ cases) and is uncorrelated with the other two factors, how well would you expect a linear model based solely on the three additive main effects could predict a criterion composed purely of the
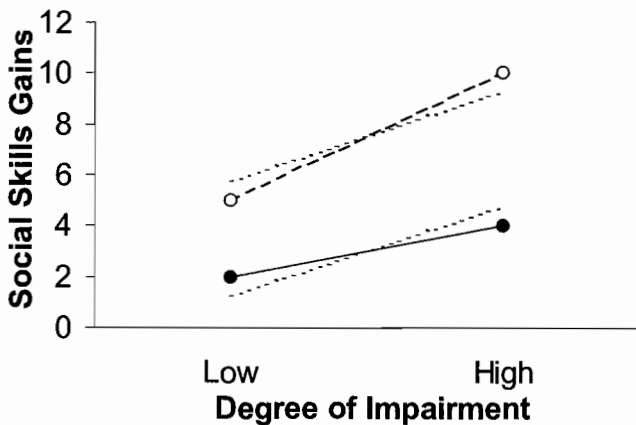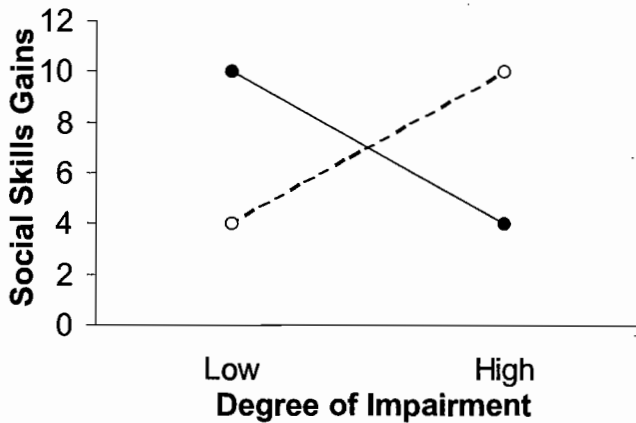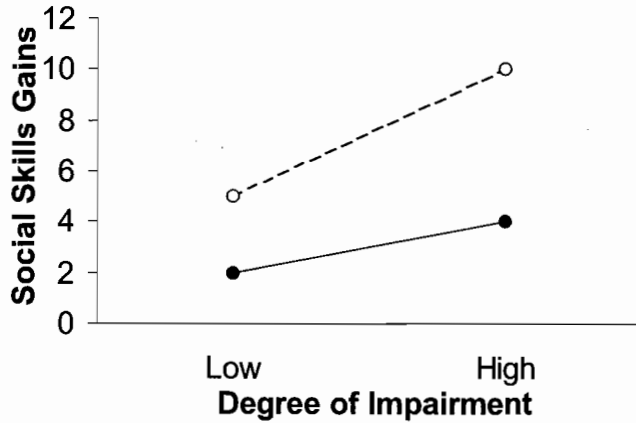
Figure. Hypothetical representations of an interaction between personality (solid lines represent introverts, dashed lines represent extraverts) and degree of impairment in predicting treatment outcome. The top panel depicts a monotonic interaction, the middle panel depicts a crossover interaction, and the bottom panel illustrates the power with which main effects approximate the monotonic interaction.

three two-way interactive relationships, $(i \times j) + (i \times k) + (j \times k)$? In this case, the multiple correlation coefficient is .97, which accounts for 94% of the variance. Thus, when variables interact monotonically, one can obtain surprisingly valid predictions even when completely ignoring interactions.

Dawes (1979) suggested that monotonic interactions are the norm, and this assertion has not been empirically challenged. The significance of monotonicity is that simple linear models based on additive main effects generate predictions of impressive accuracy. The psychological impossibility of taking into account a large number of interacting variables undermines the feasibility of holism, and the predictive power of even noninteractive judgments suggests that holistic judgment may often be a misdirected goal.

THE ETHICS OF HOLISM:
ALTERNATIVES AND MORE MODERATE FORMS

In a sense, the emerging picture of holism can be seen as a scientist's curse and a charlatan's dream. Holists imply an ability to take into account all relevant factors, which they cannot; shield their predictive failures with linguistic evasions; and often refuse to state what factors are relevant to their judgments and decisions or assess and integrate these variables in a consistent way. Such a lack of standards or guidelines based on empirical support would be debilitating to a conscientious scientist-practitioner. By appealing to vague, unsystematic processes, holists may enjoy a freedom from the extremely challenging work of developing and validating a set of guidelines for reaching important judgments and decisions, learning and teaching these guidelines to others, and constraining their practice in accordance with established professional standards. Thus, holists' claims to provide individuals seeking professional advice with better results than can well-trained scientist-practitioners is questionable on ethical grounds. Professional judgments should be based on the best available method, not unsubstantiated promises (Dawes, 1994; Faust & Ziskin, 1988).

Meehl (1986) attributed what he refers to as irrational adherence to an inferior decision-making procedure to several sources. Many individuals' unwavering belief in the efficacy of their own judgment or in the importance of their preferred theoretical identification (as contrasted, for example, with an atheoretical SPR) is a potent stumbling block. Perhaps most destructive of all is the common complaint that the use of SPRs can feel "dehumanizing," that they somehow deny the unique-

ness of individuals. This is simply untrue. In fact, research comparing clinical and statistical prediction has involved consideration of *identical* information; the only issue is how best to combine it. The importance of the feel of a procedure pales in significance when compared with a more ethically defensible benchmark for evaluating decisions: a demonstrable track record of empirical accuracy. Meehl (1986) put it this way:

> If I try to forecast something important about a college student, or a criminal, or a depressed patient by inefficient rather than efficient means, meanwhile charging this person or the taxpayer 10 times as much money as I would need to achieve greater predictive accuracy, that is not a sound ethical practice. That it feels better, warmer, and cuddlier to me as predictor is a shabby excuse indeed (p. 374).

In what might be an effort to make ethically questionable practices appear palatable, holists disparage scientific theory and method through the use of ambiguous language. This weakness can be demonstrated through a pair of well-defined problems for which the relevant information is easily ascertained and simply processed. First, if you broke your leg, would you prefer to have a doctor put it in a cast and give you crutches, or have a holistic medical practitioner treat you as a *whole person* by giving you a thorough physical examination and interpreting the significance of your broken leg only in the context of all other aspects of your physical health? A holistic healer might prescribe some mixture of acupuncture, herbs, homeopathy, magnets, nutritional supplements, reflexology, Therapeutic Touch, and other alleged remedies. It seems likely that even the most staunch advocate of holism, upon breaking his or her leg, would recognize the validity of an X ray and prefer that her physician *not* take into account other factors by simply applying a cast and prescribing some crutches. Second, it is telling that none of the American citizens who contracted anthrax in the fall of 2001 opted for a holistic assessment or remedy (Park, 2001). The whole person is not affected by the spores of this disease, other factors are not relevant to making informed decisions, and the antibacterial treatments of scientific medicine are clearly indicated.

For many, the value of clearly defining the relevant aspects of a problem becomes more murky when thinking about less directly observable events. But why should it? When your car breaks down, do you want a mechanic to assess and repair the specific mechanical failure or to holistically assess and repair the *whole car*? If this choice seems simple enough to you, you might be surprised by a question posed on the American Holistic

Veterinary Medical Association's (AHVMA) Web site, which aggressively promotes a wide range of holistic healing modalities for pets. After noting that "conventional" treatments "are employed simply to make the symptoms go away," the following query reveals the holist's hostility to science: "Picture a car with a low oil warning light. Extinguishing the light will certainly make the sign go away, but will it solve the problem?" The clear implication is that scientific veterinarians are so inept that they would extinguish the light in some way *other* than adding oil.

In reality, this scenario misrepresents the source of disagreement between the scientific and holistic approaches to professional practice: What should be done to address a well-defined problem? A scientist would assess the potentially relevant features of the problem in order to devise a sound solution, whereas a holist would view the apparent problem as indicative of a larger one and integrate a wide range of information, some relevant and some not. But in this example, the whole car is not broken—so what, one might ask, would a holistic auto mechanic do? (The holistic vets are silent on that point.) Thus, even the example chosen to disparage science shows that it involves the smartest course of action: to hone in on the specific problem (low oil) and solve it using the method suggested with the greatest likelihood of success (adding oil).[1]

In clinical practice, real problems will seldom be as simple to identify as a broken leg or low oil. Diagnostic co-occurrence, for example, is the rule rather than the exception, and a wide range of factors are potentially relevant to an understanding of causes, course, and successful treatment. There may be multiple problems and multiple goals; however, the complexity of the task, in and of itself, is poor justification for using an unnecessarily complex judgment process. One must sort the relevant from irrelevant information and combine it to reach sound decisions, and there is no evidence to suggest that a holistic reasoning strategy will be helpful.

For all practical purposes, the whole chart of astrology, the whole person of holistic health care, and the whole car of holistic auto mechanics are convenient fictions. Although they may lend a superficial plausibility to holistic practitioners' claims due to representative thinking—"a complex problem requires a complex solution" is merely a "like goes with like" assertion—they contribute

---

1. Of course, if the low oil warning soon returns, this suggests that there is in fact a larger problem in need of attention. In the first instance, simply adding oil would be the solution of choice for anyone whose philosophy of automotive repair is guided by Occam's razor, following the base rates, or any other reasonable strategy.

little or nothing to an understanding of reality. Through the use of evasive language, holists sidestep the burdens of determining what information is relevant and following a justifiable procedure for assessing and integrating this information to reach a decision. Anyone who turns his or her back on predictive methods that achieve greater empirical accuracy is arguably acting unethically.

As noted earlier, several methods have repeatedly been shown to be superior to unaided human judgment for the integration of clinical information, such as SPRs or "bootstrapped" formulas derived from analyses of expert judgment (Hoffman, 1960; Ruscio, 1998a). Indeed, one need not even consider the accuracy levels that SPRs would achieve with interaction terms included when empirically justified, for holistic judgments may often be outstripped even by purely linear SPRs. Unfortunately, SPRs are not presently available for many clinical tasks. There are, however, alternative strategies for decision making that take advantage of some of the desirable features of SPRs. Indeed, an increasingly actuarial approach is being used in several critically important domains of clinical practice. For example, state-of-the-art methods informing assessment, prevention, and treatment in areas such as child welfare (Ruscio, 1998a; Sicoly, 1989), sex offenders (Knight, 1999; Knight & Cerce, 1999), suicide (Jobes, Jacoby, Cimbolic, & Hustead, 1997; Rudd, Joiner, & Rajab, 2001), and violent behavior (Douglas, Ogloff, Nicholls, & Grant, 1999; Gardner, Lidz, Mulvey, & Shaw, 1996; Monahan et al., 2000) are all based on core principles of statistical decision making.

Though the relevant predictors and outcomes differ markedly across these domains, there is a common focus on determining which information is relevant to professional judgment and decision making, assessing it in a standardized fashion, and combining it in a simple, additive manner (often summing items on a risk inventory). The use of rating forms or checklists steers clinicians toward a consideration of the information that prior research has shown to be most pertinent. Supplementary information must be sufficiently compelling to countervail the findings of an actuarial assessment, and clinicians who identify "broken leg" counterexamples (Meehl, 1954) with due caution will achieve greater hit rates in the long run. Moreover, reaching decisions on the basis of an additive combination of relevant data increases one's consistency and accuracy. Thus, even in the absence of a SPR, one need not resort to holistic judgment.

All things considered, the ethical foundation of a holistic approach to judgment and decision making in clinical practice is highly questionable. Moreover, by endorsing practices likely to lower the net welfare of clients, those who *advocate* a holistic approach to professional practice are also acting unethically (Ruscio, 2002). Serious logical and psychological challenges are not adequately addressed by supportive research, and viable alternative methods achieve superior utility. It is telling that the premiere issue of the American Psychological Society's new journal *Psychological Science in the Public Interest* (Swets, Dawes, & Monahan, 2000) was devoted to such alternative methods for improving decision making. Enduring beliefs in the efficacy of holistic judgment may constitute one of the most significant impediments to the implementation of tried-and-tested, empirically based techniques that can improve human welfare not only in the clinic, but in a tremendous number of important domains such as finance, health, productivity, and safety.

CONCLUSIONS

In the strong form treated here, holistic judgment requires the consideration of an interaction effect of the highest order. Because of this, it suffers from the shortcomings of the clinical approach to decision making. Unaided human judgment is incapable of dealing effectively with large amounts of complex information, for our ability to identify the relevant data and combine it well is limited. Furthermore, there is an overstated need for holistic judgment in the first place. A simple, mechanical combination of a handful of relevant variables is often sufficient to achieve the predictive validity afforded by extant knowledge.

However, might it be the case that more limited forms of holism are useful? There may be several ways in which a relaxation of the extreme version of holistic judgment could prove fruitful. First, experts may be able to reason in terms of lower-order interaction effects, for the cognitive requirements would become much more manageable. Second, interactions that involve one or more categorical variables—ideally, dichotomous ones—are also less cognitively demanding than those between exclusively continuous variables. For example, experts probably think in terms of categorical relations such as "test X is valid only for population Y" quite frequently and naturally. Third, it may be easier to process an interactive relationship when it is derived from a well-supported causal theory than from a probabilistic association. Particularly at the stage of intervention, causal relationships are particularly important source of clinical decision making. Though speculative, these qualifications suggest the potential utility of appropriately con-

trained applications of holistic judgment. By backing off from the claim that "everything influences everything else," conditions in which there are a few categorical and causally important variables may support a limited form of holistic judgment. However, just as research shows that clinicians overidentify "broken leg" counterexamples to actuarial predictions (Dawes, Faust, & Meehl, 1989; Grove et al., 2000), proponents of holistic judgment should be careful not to overidentify situations that warrant it. Moreover, proponents should choose instances conducive to holism in accordance with empirically validated interaction effects, rather than proceeding based on nothing more than presumed and often unspecified interactions.

It is noteworthy that clinical judgments—and even less so holistic claims—are rarely tolerated when large sums of money are at stake. For example, when making decisions involving loans, insurance rates, or gambling odds, actuarial decision making is the norm. Within clinical psychology, the move toward SPRs has been unjustifiably slow and ineffectual (e.g., Hilton & Simmons, 2001). In the few cases in which SPRs are being actively developed, tested, or implemented, the motivation often stems from financial accountability, legal defense, or another source of external pressure. For example, by using an actuarial risk assessment process, a professional can demonstrate to a judge or jury that she used the most valid technique presently available when a decision turns out to have negative consequences.

We should insist on no less in all important realms of human affairs, where professional judgments and decisions can have a profound impact on people's lives and well-being. Holistic practitioners allege that one must consider the whole person to make judgments and reach decisions, but it is a poor reflection on mental health practices whenever professionals do not insist upon the best available methods until threatened with legal actions or financial recriminations for using suboptimal methods. Clinicians have a responsibility to their clients to scrutinize the logic and empirical evidence underlying an appeal to any approach to decision making and to adopt the most ethically defensible approach to professional practice, that which achieves the greatest observable track record of success.

REFERENCES

Chapman, L. J., & Chapman, J. (1967). Genesis of popular but erroneous diagnostic observations. *Journal of Abnormal Psychology, 72,* 193–204.

Chapman, L. J., & Chapman, J. (1971, November). Test results are what you think they are. *Psychology Today,* 18–22, 106–110.

Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications.* San Diego: Academic Press.

Dawes, R. M. (1964). Social selection based on multidimensional criteria. *Journal of Abnormal and Social Psychology, 68,* 104–109.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34,* 571–582.

Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth.* New York: Free Press.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81,* 95–106.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668–1674.

Dawes, R. M., Faust, D., & Meehl, P. E. (1993). Statistical prediction versus clinical prediction: Improving what works. In G. Keren & C. Lewis (Eds.), *Handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 351–367). Hillsdale, NJ: Erlbaum.

Dean, G., Mather, A., & Kelly, I. W. (1996). Astrology. In G. Stein (Ed.), *Encyclopedia of the paranormal* (pp. 47–99). Buffalo, NY: Prometheus.

Doane, D. C. (1956). *Astrology: 30 years research.* Hollywood, CA: Professional Astrologers.

Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L., & Grant, I. (2000). Assessing risk for violence among psychiatric patients: The HCR-20 Violence Risk Assessment Scheme and the Psychopathy Checklist: Screeing Version. *Journal of Consulting and Clinical Psychology, 67,* 917–930.

Einhorn, H. J. (1971). Use of nonlinear, noncompensatory models as a function of task and amount of information. *Organizational Behavior and Human Performance, 6,* 1–27.

Faust, D., & Ziskin, J. (1988). The expert witness in psychology and psychiatry. *Science, 241,* 31–35.

Ganzach, Y. (1994). Theory and configurality in expert and layperson judgment. *Journal of Applied Psychology, 79,* 439–448.

Ganzach, Y. (1995). Nonlinear models of clinical judgment: Meehl's data revisited. *Psychological Bulletin, 118,* 422–429.

Ganzach, Y. (2001). Nonlinear models of clinical judgment: Communal nonlinearity and nonlinear accuracy. *Psychological Science, 12,* 403–407.

Gardner, W., Lidz, C. W., Mulvey, E. P., & Shaw, E. C. (1996). Clinical versus actuarial predictions of violence in patients with mental illnesses. *Journal of Consulting and Clinical Psychology, 64,* 602–609.

Gilovich, T., & Savitsky, K. (1996). Like goes with like: The role of representativeness in erroneous and pseudoscientific beliefs. *Skeptical Inquirer, 20,* 34–40.

Goldberg, L. R. (1968). Simple models or simple processes? Some research on clinical judgments. *American Psychologist, 23,* 483–496.

Goldberg, L. R. (1991). Human mind versus regression equation: Five contrasts. In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology* (Vol. 1, pp. 173–184). Minneapolis: University of Minnesota Press.

Graham, J. R. (2000). *MMPI-2: Assessing personality and psychopathology* (3rd ed.). New York: Oxford University Press.

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

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12,* 19–30.

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): I. Construction of the schedule. *Journal of Psychology, 10,* 249–254.

Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behavior, 25,* 393–408.

Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psychological Review, 57,* 116–131.

Jobes, D. A., Jacoby, A. M., Cimbolic, P., & Hustead, L. A. T. (1997). Assessment and treatment of suicidal clients in a university counseling center. *Journal of Counseling Psychology, 44,* 368–377.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

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

Kelly, I. W. (1997). Modern astrology: A critique. *Psychological Reports, 81,* 1035–1066.

Kelly, I. W. (1998). Why astrology doesn't work. *Psychological Reports, 82,* 527–546.

Knight, R. A. (1999). Validation of a typology for rapists. *Journal of Interpersonal Violence, 14,* 303–330.

Knight, R. A., & Cerce, D. D. (1999). Validation and revision of the Multidimensional Assessment of Sex and Aggression. *Psychologica Belgica, 39,* 135–161.

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.

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest, 1,* 27–66.

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

Meehl, P. E. (1967). What can the clinician do well? In D. N. Jackson & S. Messick (Eds.), *Problems in human assessment.* New York: McGraw-Hill.

Meehl, P. E. (1986). Causes and effects of my disturbing little book. *Journal of Personality Assessment, 50,* 370–375.

Meehl, P. E. (1993). When shall we use our heads instead of the formula? *Journal of Counseling Psychology, 4,* 81–89.

Merlo, L., & Barnett, D. (2001). All about inkblots. *Scientific American, 285,* 13.

Monahan, J., Steadman, H. J., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., Silver, E., Roth, L. H., & Grisso, T. (2000). Developing a clinically useful actuarial tool for assessing violence risk. *British Journal of Psychiatry, 176,* 312–319.

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

Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology, 13,* 248–277.

Park, R. (2001, October 19). Bioterrorism: So far, the count is one dead. *What's New.*

Rudd, M. D., Joiner, T. E., & Rajab, M. H. (2001). *Treating suicidal behavior: An effective, time-limited approach.* New York: Guilford.

Ruscio, J. (1998a). Information integration in child welfare cases: An introduction to statistical decision making. *Child Maltreatment, 3,* 143–156.

Ruscio, J. (1998b). The perils of post-hockery. *Skeptical Inquirer, 22,* 44–48.

Ruscio, J. (2000a). Risky business: Vividness, availability, and the media paradox. *Skeptical Inquirer, 24,* 22–26.

Ruscio, J. (2000b). The role of complex thought in clinical prediction: Social accountability and the need for cognition. *Journal of Consulting and Clinical Psychology, 68,* 145–154.

Ruscio, J. (2002). *Clear thinking with psychology: Separating sense from nonsense.* Pacific Grove, CA: Wadsworth.

Ruscio, J., & Stern, A. (2001). The consistency and accuracy of holistic judgment: Clinical decision making with a minimally complex task. Manuscript submitted for publication.

Sicoly, F. (1989). Prediction and decision making in child welfare. *Computers in Human Services, 5,* 43–56.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance, 6,* 649–744.

Stewart, T. R. (1988). Judgment analysis: Procedures. In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 41–74). Amsterdam: North-Holland Elsevier.

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

Turk, D. C., & Salovey, P. (1988). *Reasoning, inference, and judgment in clinical psychology.* New York: Free Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic

for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

Van Rooij, J. (1994). The whole chart and nothing but the whole chart. *Correlation, 13*, 54–56.

Wood, J. M., Lilienfeld, S. O., Nezworski, M. T., & Garb, H. N. (2001). Coming to grips with negative evidence for the Comprehensive System for the Rorschach: A comment on Gacono, Loving, and Bodholdt; Ganellen; and Bornstein. *Journal of Personality Assessment, 77*, 48–70.

Yntema, D. B., & Torgerson, W. S. (1961). Man-computer cooperation in decisions requiring common sense. *IRE Transactions of the Professional Group on Human Factors in Electronics, HFE-2*, 20–26.