Research Report

PSYCHOLOGICAL MODELS OF PROFESSIONAL DECISION MAKING

Mandeep K. Dhami

City University, London, United Kingdom, and University of Victoria, Victoria, British Columbia, Canada

Abstract—*People are often expected to make decisions based on all of the relevant information, weighted and combined appropriately. Under many conditions, however, people use heuristic strategies that depart from this ideal. I tested the ability of two models to predict bail decisions made by judges in two courts. In both courts, a simple heuristic proved to be a better predictor of judicial decisions than a more complex model that instantiated the principles of due process. Specifically, judges were "passing the buck" because they relied on decisions made by the police, prosecution, and previous bench. Problematically, these earlier decisions were not significantly related to case characteristics. These findings have implications for the types of models researchers use to capture professional decision-making policies.*

Ideally, we expect decision makers to use all of the relevant information, and weight and combine it appropriately. Moreover, we expect them to behave like this when their decisions have significant consequences. For more than 50 years, researchers have captured judgment policies in domains such as medicine (see Wigton, 1996), education (see Heald, 1991), and accounting (see Waller, 1988) using multiple linear regression. This model depicts professionals as behaving in an ideal way. It is reported that people combine multiple differentially weighted cues in a compensatory way, so, for example, a low weight attached to one cue is compensated by a high weight attached to another cue. However, the regression approach assumes large attentional, memory, and processing abilities, and ignores the impact of sequential processing (e.g., Dhami & Harries, 2001; Gigerenzer, Todd, & the ABC Research Group, 1999). This approach is also inflexible because it assumes the same cues are used to make decisions on different cases. Furthermore, policy-capturing researchers have overlooked the fact that decision strategies are adapted to the demands of the task. For instance, under conditions of time pressure, people tend to use fewer cues and simple noncompensatory strategies, so, for example, an initial leaning toward a decision based on a cue with a high weight will not be altered by cues with lower weights (e.g., Payne, Bettman, & Johnson, 1993; Rieskamp & Hoffrage, 1999).[1]

There are several nonstatistical and cognitively simpler strategies that represent viable alternatives to the regression model. Two such models are Franklin's rule and the matching heuristic. The processes by which these models predict whether a judge makes a punitive bail decision[2] are described in the appendices. Like the regression model, Franklin's rule (originally described by Benjamin Franklin) involves the compensatory combination of multiple differentially weighted cues, and is limited in its inflexible cue use. However, it differs from the regression model in that it does not compute optimal weights as in the least squares regression, nor does it take into account the interdependencies among cues. By contrast, the matching heuristic (Dhami & Ayton, 1998, 2001) uses an even simpler cue-weighting method, searches through a small subset of the cues, and bases its predictions on one cue alone. It is noncompensatory because a decision is based on the value of one cue, and so is not altered by values of other cues. It is also flexible because different cues can be used to make decisions on different cases. The matching heuristic is therefore a "simple" or "fast and frugal" heuristic (see Gigerenzer et al., 1999).

To date, most research comparing the predictive validity of the regression model and these simple heuristics has been based on simulations in which models predict a criterion. Studies show that whereas the regression model is the best predictor of a criterion at the model-fitting stage, simple heuristics tend to outperform the regression model at the cross-validation stage (see Gigerenzer & Goldstein, 1996; Gigerenzer et al., 1999).[3] In behavioral studies, the matching heuristic performed as well as the regression model when predicting doctors' prescription decisions (Dhami & Harries, 2001), and outperformed Franklin's rule when predicting judges' bail decisions (Dhami & Ayton, 2001). In both studies, however, participants made decisions on systematically designed hypothetical cases (in which cues are independent), which are common in policy-capturing research. The validity of the captured policies is thus questionable (e.g., Ebbesen & Konecni, 1975; Phelps & Shanteau, 1978). Furthermore, critics argue that support for simple heuristics is needed from behavioral data gathered under naturalistic conditions (e.g., Lipshitz, 2000). Indeed, some may consider that a strong test of these heuristics would involve participants, such as judges, who are explicitly guided and motivated to reason in a manner that is neither fast nor frugal.

THE PRESENT STUDY

Like most professional decisions, judicial decisions are guided by formal rules. In Anglo-American jurisdictions, judicial decisions must comply with the principles of due process. In theory, when deciding to convict, judges or jurors should search through all information pertaining to guilt and innocence, weight it according to its reliability and validity,

The research presented in this article received the 2001–2002 Social Issues Dissertation Award, first prize, from the Society for the Psychological Study of Social Issues, Division 9 of the American Psychological Association. Address correspondence to Mandeep K. Dhami, Department of Psychology, University of Victoria, P.O. Box 3050, STN CSC, Victoria, B.C., V8W 3P5, Canada; e-mail: mkdhami@uvic.ca.

1. Indirect support for the use of simple strategies is obtained from studies reporting only one statistically significant beta weight in the regression model (e.g., Deshpande & Schoderbek, 1993; Gonzalez-Vallejo, Sorum, Stewart, Chessare, & Mumpower, 1998).

2. In the United Kingdom, remanding a defendant into custody or granting bail subject to conditions is referred to as a punitive decision, and releasing a defendant on unconditional bail is a nonpunitive decision.

3. At the model-fitting stage, the model is used to make predictions on the cases that were used to construct the model. At the cross-validation stage, the model is used to make predictions on a new, equally sized sample of cases.

Table 1. *Observed cues and their values*

Cue	Values
Defendant's characteristics	
Age	18–20/21+
Gender	Male/female
Race	White/visible ethnic group
Strength of community ties	Has job or child or partner or home/has none of these
Current offense	
Seriousness of offense	Trial in lower courts/trial by jury in higher court
Category of offense	Against person/against property or other
Number of offenses	1/2+
Victim	Known or unknown person(s)/consensual crime or business victim
Is defendant solely involved?	Yes/no
Plea	Guilty/not guilty/no plea
Strength of prosecution case	Has physical evidence or witness/has none of these
Maximum penalty if convicted	Custodial/noncustodial
Defendant's previous record	
Previous convictions	None/yes-similar/yes-dissimilar
Bail record	None or good/breached bail
Bail hearing	
Is defendant in court?	Yes/no
Is defendant legally represented?	No/yes by own or court-appointed solicitor
Who is the prosecutor?	Crown prosecution service/other
Circumstances of adjournment	For trial/for sentence or appeal or other reason
Who requested the adjournment?	Defense/prosecution/court
Length of adjournment	1 week/2 weeks/3 weeks/4 weeks/ 5 weeks/6 weeks
Number of previous adjournments	0–1/2+
Prosecution request	Do not oppose bail/ask for conditions or oppose bail
Defense request	Unconditional bail/conditional bail/no application for bail
Previous court bail decision	None/unconditional bail/conditional bail or remand in custody
Police bail decision	Unconditional bail/conditional bail or remand in custody

Note. Values separated by "or" were observed separately, but were combined for analysis.

and combine it, so that, for example, an initial leaning toward a verdict of guilt can be altered by evidence indicating innocence (Packer, 1968). A similar process is advocated for making bail decisions. Like most judicial decisions, bail decisions have huge ramifications for defendants and the public. The bail decision is one of the most frequent decisions made by judges, and may influence later decisions to convict and sentence (Davies, 1971). The present study compared the ability of Franklin's rule and the matching heuristic to predict the bail decisions made by judges on real cases appearing in real time. On the basis of past psychological research, I hypothesized that the matching heuristic would be the better predictor of judges' decisions. By contrast, a hypothesis derived from legal theory (that judges would observe the principles of due process) suggests that Franklin's rule will outperform the simple heuristic in predicting judicial decisions. This study pitted these two hypotheses against one another.

METHOD

Observers and Observed Judges

The decisions made by benches of judges in two London, United Kingdom, courts were observed over a 4-month period. Observer 1 recorded 159 decisions made by 25 benches in Court A, and Observer 2 recorded 183 decisions made by 32 benches in Court B. The benches comprised different combinations of 55 judges in Court A and 56 judges in Court B. There was no significant difference between the average years of experience of judges sitting in Court A ($M = 10.1$, $SD = 7.8$) and Court B ($M = 9.5$, $SD = 7.3$), $t(108) = 0.59$.[4]

4. All tests are two-tailed.

Observational Coding Scheme

Details of the cases presented and the decisions made were recorded using a structured coding scheme. Construction of the scheme was informed by a task analysis, and the scheme was pilot-tested on 15 bail hearings observed in 1 week in a third court. Data were recorded on 25 verbal, nonverbal, and written cues that the task analysis indicated may be available to judges during bail hearings. The cues are shown in Table 1. They can be divided into those referring to (a) the personal characteristics of the defendant, (b) the offense with which the defendant is charged, (c) the defendant's previous record, and (d) the bail hearing. In addition to recording details of each case and the decision, observers measured the duration of bail hearings using a stopwatch.

Interobserver reliability

Interobserver reliability was assessed in the middle of the observation period, when both observers recorded data on 26 hearings in 1 week in the two courts (i.e., 8 in Court A and 18 in Court B). Calculation of Cohen's kappa indicated that agreement ranged from perfect (i.e., 1.0) to excellent (i.e., $\geq$.75) on most variables. The recorded duration of bail hearings was also highly consistent between the two observers, $r = .98$, $p < .001$.

Observed cases: Availability and intercorrelations of cues

Information was often unavailable to judges for the following 4 of the 25 cues: defendant's previous convictions, defendant's bail record, defendant's community ties, and bail decision by the police. Chi-square analyses revealed that compared with Court A, in Court B, a greater percentage of defendants were present in the courtroom during the hearing, were of ethnic origin, were legally represented, and had been charged with crimes against a person, and a smaller percentage had pleaded guilty and had previous adjournments ($p < .05$). First-order intercorrelations among the 25 cues were computed for each court. Seventy-three of the coefficients in Court A and 58 in Court B were statistically significant ($p < .05$), although none would be if a Bonferroni correction were applied. The mean cue intercorrelation was .2 ($SD = .3$) in Court A and .1 ($SD = .3$) in Court B.

RESULTS

Bail Hearings: Decisions and Duration

There was a significant difference between the proportion of punitive decisions made in Court A (i.e., 40.9%) and Court B (i.e., 54.1%), $\chi^2(2, N = 342) = 7.76$, $p < .05$. Furthermore, the duration of bail hearings in Court A ($M = 6.7$ min, $SD = 6.0$) was significantly different from the duration in Court B ($M = 9.5$ min, $SD = 8.4$), $t(312) = 3.54$, $p < .05$.

Bail Decision-Making Policies

Franklin's rule and the matching heuristic were used to capture the policies of each court separately because the courts differed in the cases presented, the decisions made, and the duration of hearings. Policies were not captured for individual benches because benches are not stable groups—judges are constantly rotated, and individual benches make too few decisions for meaningful analysis. The 25 cues were simplified (most converted to binary cues) for ease of analysis (see Table 1). This process was informed by the task analysis and was compatible with past research (Dhami & Ayton, 2001). As in past research, both models were constructed so that they aimed to predict a punitive decision and made nonpunitive decisions only by default. The models were formed so that they treated unavailable cue information in a similar way, and the two models computed the same number of parameters. Each court's decisions were randomly divided into a modeling set and a cross-validation set (i.e., 80 modeling and 79 cross-validation cases in Court A, and 92 modeling and 91 cross-validation cases in Court B). So that an idiosyncratic division would be avoided, this process was repeated 10 times, yielding 10 different modeling and cross-validation sets for each court. Each time the model was constructed on the modeling set, and predicted decisions first for this set and then for the cross-validation set.

Whereas Franklin's rule searched through all 25 cues, the maximum number of cues searched (K) by the matching heuristic was on average 3.0 ($SD = 0.7$) for Court A and 2.8 ($SD = 0.4$) for Court B. As Table 2 shows, despite this large difference in cue use between the models, for both courts, the matching heuristic outperformed Franklin's rule when predicting decisions at the model-fitting stage. Although the predictive power of both models was reduced at the cross-validation stage, the matching heuristic remained the better predictor of judges' decisions for both courts. A similar pattern of results emerged in a comparison of the ability of the models to predict the nonpunitive and punitive decisions separately (with the exception that Franklin's rule outperformed the matching heuristic in predicting Court B's nonpunitive decisions).

The maximum number of cues searched (K) by the matching heuristic and the rank order of cues differed slightly across the 10 tests because the properties of the modeling set changed from test to test. For illustrative purposes, Figure 1 shows the matching heuristic for each court, where K and the percentage of correct predictions was close to the mean found at the model-fitting stage. The model in Figure 1a correctly predicted 96.3% of decisions in Court A. The model in Figure 1b correctly predicted 94.6% of decisions in Court B.

DISCUSSION

In the present study, judicial decisions made in two courts were better predicted by the matching heuristic than by Franklin's rule. The matching heuristic depicts judges as basing decisions on one cue. Judges' reliance on the decisions made by the police, previous bench, and prosecutor (see Figs. 1a and 1b) suggests that they were either intentionally or unintentionally "passing the buck." (Note that these cues were not significantly correlated with other cues such as the nature and seriousness of the offense.) Although this study does not bear upon the accuracy of the decisions, judges behaved contrary to the ideals of due process, according to which the number of innocent defendants who are treated punitively should be minimized. Converging evidence for the fast and frugal nature of judicial decisions derives from the observed brevity of the bail hearings and the consequent rapidity with which decisions must have been made.

The present findings support the validity of simple heuristics in capturing decision policies under naturalistic conditions and in the group context. In fact, the predictive validity of the matching heuristic was greater than that reported in the past (Dhami & Ayton, 2001; Dhami & Harries, 2001), and greater than the predictive validity of other simple heuristics (see Gigerenzer et al., 1999) and the regression model (see Brehmer & Brehmer, 1988).

Several conditions may have enabled the heuristic strategy to prevail in the present study. First, judges were presented with numerous cues and were often faced with a heavy caseload. There is evidence that people switch to simple noncompensatory strategies that use few cues as the number of cues increases and as time pressure increases (e.g., Payne et al., 1993; Rieskamp & Hoffrage, 1999). Second, judges made decisions as a bench. Groups making decisions involving shared responsibility tend to use few cues (Weldon & Gargano, 1985). Finally, the law affords judges considerable discretion concerning the cues they use to make their decisions. Notably, however, these conditions are not very dissimilar from those faced by professionals making decisions in other domains.

The present findings have implications for policy-capturing research. One explanation for the popularity of the regression model in policy-capturing research relates to the efficacy of alternative models available in the past. For instance, the conjunctive and disjunctive models (Einhorn, 1970) have performed poorly relative to the regression model (Ogilvie & Schmitt, 1979), and the predictive validity of models developed via process tracing is difficult to test (see Juslin & Montgomery, 1999). Predictive validity has been the main reason for employing the regression model (Hoffman, 1960; Stewart, 1988). As the present study demonstrates, noncompensatory simple heuristics can be excellent predictors of decisions—even outperforming compensatory models that share key features with the regression model.

Clearly, another criterion for choosing models should be their psychological plausibility (Gigerenzer et al., 1999). Simple heuristics are grounded in research on human cognitive capacities. For example, the matching heuristic uses frequencies when determining the critical value on a cue. It is claimed that this is a natural form of processing (e.g., Cosmides & Tooby, 1996). Definition of the critical value and utilization validity is supported by evidence that people use a subset of the available information when learning about the relations between cues and an outcome (Nisbett & Ross, 1980). The critical value embodies a type of positive-test bias in which only the information that indicates a focal (in this case punitive) decision is searched and used. There is general evidence for such strategies in other domains (e.g., Klayman & Ha, 1987). Finally, the heuristic embodies the idea of matching characteristics of individual cases with those of a prototype and is thus consistent with exemplar models in categorization, although they tend to be more complex (see Estes, 1994). Therefore, simple heuristics also meet the criterion of psychological plausibility.

A sound psychological theory of human decision making is possible only if we test the relative predictive validity of cognitively plausible models. Future research should involve models that have been constructed so that individual decision processes (e.g., compensation, linearity, cue weighting) are systematically manipulated, and tested under different task conditions (e.g., time pressure, number of available cues, cue redundancy). A full understanding of decision processes is essential as practitioners and policymakers often rely on our help in developing appropriate tools for training, evaluating, and aiding professional decision making.

Acknowledgments—I am grateful to Michelle Gates for her assistance in conducting the observations, and Torsten Morhbach for programming the models. I also thank Laura Martignon for her helpful comments on the models, David Mandel and Peter Ayton for their comments on early drafts of the manuscript, and Barbara Spellman and three anonymous reviewers.

Table 2. *Mean percentage of court decisions predicted correctly by the models*

	Model			
	Franklin's rule		Matching heuristic	
Test stage and decisions	*M*	*SD*	*M*	*SD*
Court A				
Model-fitting				
Overall	89.1	3.2	95.4	1.6
Nonpunitive	86.4	5.2	92.5	2.2
Punitive	93.5	4.3	99.5	1.7
Cross-validation				
Overall	86.3	2.7	91.8	3.6
Nonpunitive	81.1	6.4	89.0	5.1
Punitive	93.3	3.3	95.2	8.2
Court B				
Model-fitting				
Overall	82.3	3.8	91.6	2.8
Nonpunitive	87.3	2.2	86.8	28.9
Punitive	78.3	6.7	95.5	1.5
Cross-validation				
Overall	73.4	4.9	85.4	22.1
Nonpunitive	78.7	9.6	77.9	37.1
Punitive	68.8	6.8	92.9	8.8

Note. Means and standard deviations are calculated over 10 tests.

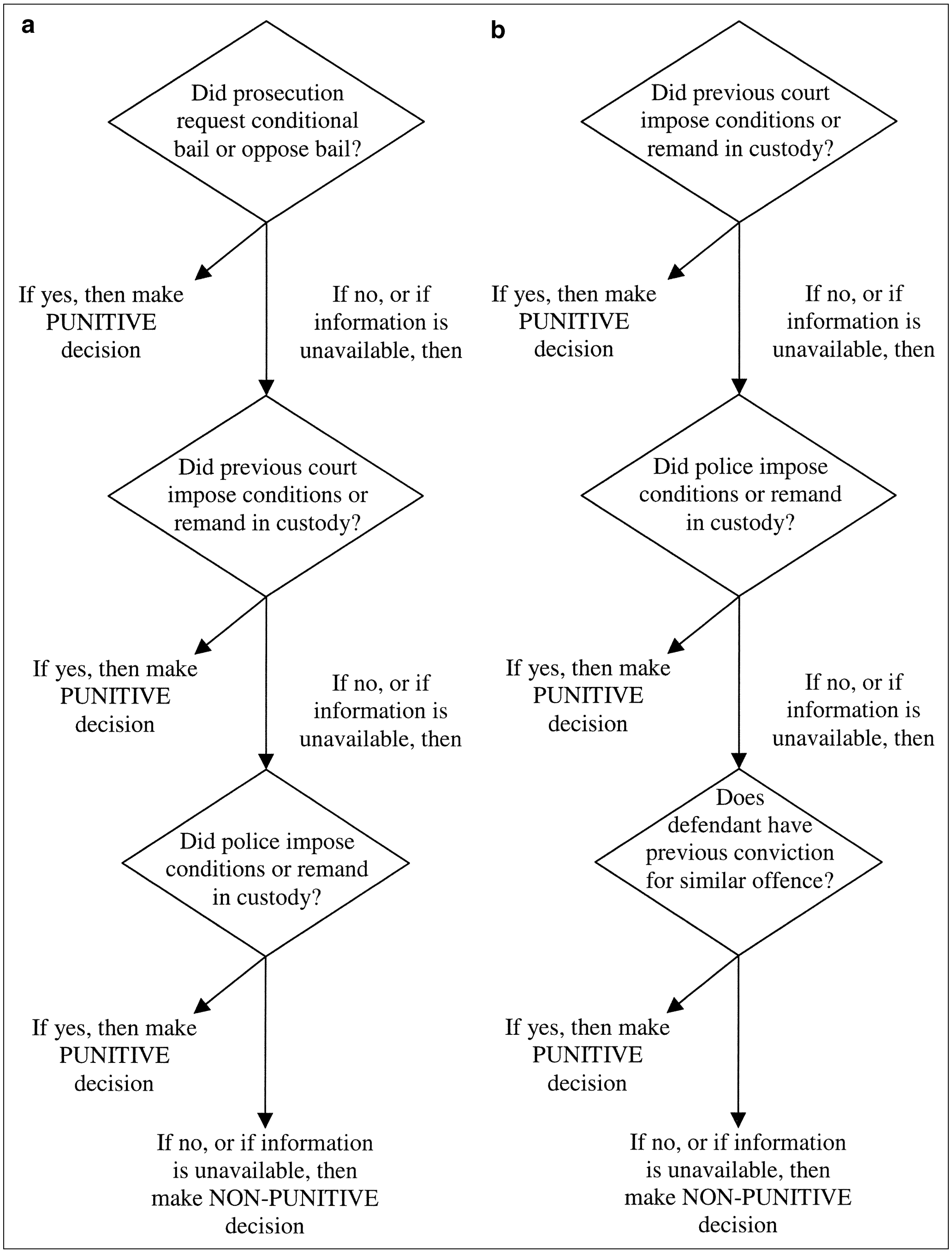

Fig. 1. Matching heuristics for Court A (a) and for Court B (b).

REFERENCES

Brehmer, A., & Brehmer, B. (1988). What have we learned about human judgment from thirty years of policy capturing? In B. Brehmer & C.R.B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 75–114). Amsterdam: North-Holland.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1–73.

Davies, C. (1971). Pre-trial imprisonment: A Liverpool study. *British Journal of Criminology*, *11*, 32–48.

Deshpande, S.P., & Schoderbek, P.P. (1993). Pay-allocations by managers: A policy-capturing approach. *Human Relations*, *46*, 465–479.

Dhami, M.K., & Ayton, P. (1998, November). *Legal decision making the fast and frugal way.* Poster presented at the annual meeting of the Society for Judgment and Decision Making, Dallas, TX.

Dhami, M.K., & Ayton, P. (2001). Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making*, *14*, 141–168.

Dhami, M.K., & Harries, C. (2001). Fast and frugal versus regression models of human judgment. *Thinking and Reasoning*, *7*, 5–27.

Ebbesen, E.B., & Konecni, V.J. (1975). Decision making and information integration in the courts: The setting of bail. *Journal of Personality and Social Psychology*, *32*, 805–821.

Einhorn, H.J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin*, *73*, 221–230.

Estes, W.K. (1994). *Classification and cognition*. New York: Oxford University Press.
Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.
Gigerenzer, G., Todd, P.M., & the ABC Research Group. (Eds.). (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
Gonzalez-Vallejo, C., Sorum, P.C., Stewart, T.R., Chessare, J.B., & Mumpower, J.L. (1998). Physicians' diagnostic judgments and treatment decisions for acute otitis media in children. *Medical Decision Making*, *18*, 149–162.
Heald, J.E. (1991). Social judgment theory: Applications to educational decision making. *Educational Administration Quarterly*, *27*, 343–357.
Hoffman, P.J. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin*, *47*, 116–131.
Juslin, P., & Montgomery, H. (Eds.). (1999). *Judgment and decision making: New-Brunswikian and process-tracing approaches*. Hillsdale, NJ: Erlbaum.
Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.
Lipshitz, R. (2000). Two cheers for bounded rationality [Commentary]. *Behavioral and Brain Sciences*, *23*, 756.
Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
Ogilvie, J.R., & Schmitt, N. (1979). Situational influences on linear and nonlinear use of information. *Organizational Behavior and Human Performance*, *23*, 292–306.
Packer, H.L. (1968). *The limits of the criminal sanction*. Stanford, CA: Stanford University Press.
Payne, J.W., Bettman, J.R., & Johnson, E.J. (1993). *The adaptive decision maker*. Cambridge, England: Cambridge University Press.
Phelps, R.H., & Shanteau, J. (1978). Livestock judges: How much information can an expert use? *Organizational Behavior and Human Performance*, *21*, 209–219.
Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics, and how can we tell? In G. Gigerenzer, P.M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 141–167). New York: Oxford University Press.
Stewart, T. (1988). Judgment analysis: Procedures. In B. Brehmer & C.R.B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 41–74). Amsterdam: North-Holland.
Waller, W.S. (1988). Brunswikian research in accounting and auditing. In B. Brehmer & C.R.B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 247–272). Amsterdam: North-Holland.
Weldon, E., & Gargano, G.M. (1985). Cognitive effort in additive task groups: The effects of shared responsibility on the quality of multiattribute judgments. *Organizational Behavior and Human Decision Processes*, *36*, 348–361.
Wigton, R.S. (1996). Social judgment theory and medical judgment. *Thinking and Reasoning*, *2*, 175–190.

(RECEIVED 2/13/02; REVISION ACCEPTED 6/12/02)

APPENDIX A: FRANKLIN'S RULE

In this model, cues are differentially weighted. For each case, cue values are multiplied by their weights and then summed. If the sum is equal to or greater than a threshold value, then a punitive decision is predicted.[5] If not, a nonpunitive decision is predicted.

Construction of the Model

Cue values are coded. For example, in the present study, females were coded as 0 and males as 1 for the gender cue. A threshold value for predicting a punitive decision is established by taking the sum of each case in the modeling set, totaling these sums, and dividing this total by the number of cases in the modeling set. The weight for each cue is determined from the modeling set by calculating for each cue value the proportion of cases treated punitively, comparing the proportions for the different cue values, and then taking the greatest proportion as the weight for the cue. For example, if the proportion of males treated punitively is .78 (i.e., 14 treated punitively out of 18) and the proportion of females treated punitively is .33 (i.e., 3 treated punitively out of 9), the weight for the gender cue is .78.

Example: Decision of Judge 1 on Case 3 (taken from Dhami & Ayton, 2001)

The threshold value for this judge was 3.52. Based on this judges' cue weights, consideration of Case 3 was as follows: gender(0)(0.72) + race(1)(0.67) + age(0)(0.67) + seriousness of the offense(1)(0.78) + prosecution request(1)(0.72) + past criminal record(0)(0.73) + strength of prosecution case(0)(0.78) + defendant's community ties(0)(0.67) + police bail decision(1)(0.67) = 2.84. The case sum was below the threshold; thus, Franklin's rule predicted a nonpunitive decision. In fact, the judge made a punitive decision on this case.

APPENDIX B: MATCHING HEURISTIC

In this model, cues are rank-ordered by their utilization validities. For each case, K cues are searched in order, for a critical value that indicates a punitive decision. If a critical value on a cue is found, search is terminated and a punitive decision is predicted. Otherwise, search continues until K cues have been searched, and if by this time no critical value has been found, a nonpunitive decision is predicted.

Construction of the Model

For each cue, the critical value indicating a punitive decision is the value of that cue that was most frequently treated punitively in the cases in the modeling set. For example, the critical value for the gender cue is male if more males than females were treated punitively (i.e., 14 males treated punitively compared with 3 females). (If these absolute frequencies are equal, the cue value with the lowest absolute frequency of cases treated nonpunitively is selected as the critical value; if the frequencies treated nonpunitively are also equal, a critical value is selected randomly.)

Cues are rank-ordered according to their utilization validity, which is defined as the proportion of cases with the critical value that were treated punitively in the modeling set. To continue the example, the validity of the gender cue would be the proportion of males treated punitively, or .78 (14 males treated punitively out of 18). A rank of 1 is assigned to the cue with the largest validity. (Cues with tied ranks are placed in order of their presentation to the judges.)

The maximum number of cues the heuristic searches (i.e., K) is determined by systematically testing the heuristic's ability to correctly predict decisions in the modeling set where $K = N$ cues, $K = N - 1$ cues, $K = N - 2$ cues, and so forth. The value of K that yields the greatest percentage of correct predictions is selected.

Example: Decision of Judge 1 on Case 3 (taken from Dhami & Ayton, 2001)

For this judge, the heuristic would search for information on only one cue—the seriousness of the alleged offense. It would predict a punitive decision if the offense was indictable (serious). In Case 3, the offense was serious, so the heuristic predicted a punitive decision. In fact, the judge did make a punitive decision.

5. For ease of exposition, Appendices A and B refer to constructing models for predicting punitive and nonpunitive decisions, but the procedures described would be the same for whatever decision is of interest (e.g., to prescribe or not prescribe a particular medication).