# COGNITIVE PROCESSES in CHOICE and DECISION BEHAVIOR

Edited by
THOMAS S. WALLSTEN

# COGNITIVE PROCESSES IN CHOICE AND DECISION BEHAVIOR

edited by
THOMAS S. WALLSTEN
*University of North Carolina at Chapel Hill*


LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1980 Hillsdale, New Jersey

# Contents

# 2 On the External Validity of Decision-Making Research: What Do We Know About Decisions in the Real World?

Ebbe B. Ebbesen
and Vladimir J. Konečni
*University of California, San Diego*

## THE ISSUE

Many current models of decision making are based on evidence obtained from laboratory experiments in which a relatively limited set of "simulated" decision problems have been used. For example, in the area of probabilistic inference, subjects are often presented with gambles differing in the amounts of money that can be won or lost and in the probabilities associated with such outcomes (e.g., Anderson & Shanteau, 1970; Kahneman & Tversky, 1979; Lichtenstein & Slovic, 1973; Payne, 1975; Slovic & Lichtenstein, 1968). Even when the choice alternatives do not involve monetary gain or loss, the decision problems are usually decomposed in such a way that probabilistic information is presented numerically rather than experientially (e.g., Kahneman & Tversky, 1973). In fact, the majority of what are considered to be important results in the decision-making area has been obtained with procedures in which the decision task was, at least to some extent, already decomposed into the dimensions that were of primary interest to the researcher (cf. Slovic & Lichtenstein, 1971, and Slovic, Fischhoff, & Lichtenstein, 1977). For example, when the interest is in comparing the role that certain key variables play in normative models with the actual effects of these variables on decisions, relevant decision tasks are not found in the real world, but rather are constructed in such a way that these key variables are presented in a decomposed form.

Even when decision-making models are applied to specific real-world decisions, much of the decision analyst's time is spent redefining the decision task facing the client so that its format conforms to the structure of laboratory simulations (Keeney, 1973, 1977). For example, in applying decision theory to the

problem of the distribution of fire engines in a city, Keeney (1973) first had to "discover" the attributes of fire fighting that the decision maker thought were relevant, then "elicit" the utility functions associated with these attributes. Probabilities were also "elicited" but from experts rather than from the decision maker.[1] In short, the decision problem was decomposed such that the choice alternatives made available to the client were presented as lists of attributes, each with an associated value and probability. This was done even though the original decision problem was described in a completely different manner.

In research that has been guided by linear rather than normative models, the decision tasks are somewhat less constrained to fit a preexisting theoretical mold. Nevertheless, the data being fit by the linear model have generally been obtained from people making decisions in what are obviously simulations of the relevant decision problems. Thus, it is typical that only some of the predictors that might be relevant in similar real-world decision tasks are included in the simulated task and, more importantly, the ones that are presented are usually in decomposed form (although there are exceptions, e.g., Einhorn, 1974; Exp. 2 in Phelps & Shanteau, 1978). That is, the decision maker is usually given a *list* of the levels of the relevant factors (not some holistic representation of the predecision situation) and is told to reach a decision (e.g., Anderson, 1974; Naylor & Wherry, 1965; Exp. 1 in Phelps & Shanteau, 1978; Slovic, 1969). For example, in response to a request for greater face validity of the decision tasks used to study stockbrokers' investment decisions (Slovic, Fleissner, & Bauman, 1972), Ebert and Kruse (1978) asked professional securities analysts to consider a large number of cues that had been constructed from the actual performance data of relevant securities. Thus, although the range of levels of the cues probably matched those usually seen by the analyst, the cues were still presented, one at a time, in the standard list format. Furthermore, although a major purpose of the study was to improve the realism of past research, the subjects were clearly told that their decisions were hypothetical and therefore had no monetary consequences. Finally, the authors did not report whether the intercorrelations of the cues (Hammond, Stewart, Brehmer, & Steinmann, 1975) used in the experimental task matched those in the real world.

---

[1]A major problem with applying some normative models to real-world decision problems is that some of the constructs in the models refer to subjective variables whose values cannot be observed directly. Often, the levels of these variables are obtained from ratings made by subjects. Therefore, all of the problems associated with "reactive" measurement (Webb, Campbell, Schwartz, & Sechrest, 1966) should be relevant to such applications. Unfortunately, very little attention seems to be paid to the fact that values (and other subjective states) may be *constructed* by the decision maker for the first time when asked about them. It is generally known in social psychology, for example, that attitudes are consistent with action only under very special circumstances (Zimbardo, Ebbesen, & Maslach, 1977). If our measures are tapping basic and stable states, why do multiple measures of the "same" state correlate so poorly (cf. Fischhoff, Slovic, & Lichtenstein, this volume)?

Even in several instances in which the decision task has involved stimuli presented in a holistic format, the subjects are told that their decisions are hypothetical and are also typically fully aware that their decisions are being evaluated (e.g., Phelps & Shanteau, 1978). Thus, the consequences of the decisions are rarely the same as those naturally occurring in the real-world task being simulated.[2]

One explanation for the overrepresentation of laboratory simulations in past research on decision making is that researchers have been primarily concerned with discovering what are thought to be basic psychological rules or processes. If one begins with the assumptions (1) that such rules exist; (2) that their number is probably small; and (3) that the different rules do not interact in any important way, then the major consideration in selecting a decision task should be that it will allow the researcher to clearly demonstrate the operation of one or more of these rules or processes. Because real-world decision making is bound to be clouded by a host of irrelevant and potentially confounding factors, constructing a decision task provides the opportunity to conduct more controlled and inexpensive research. In fact, because the results from such research are likely to reflect the operation of a "pure" process or rule, unconfounded by other factors, the conclusions that are reached about decision making on the basis of laboratory simulations should have great generality.

## THE EVIDENCE

What evidence do we have to support such a view of laboratory simulations? We would argue, little or none. In fact, what evidence there is suggests that this view might be incorrect.

### Task Specificity

Consider first the picture that is emerging from the laboratory simulations currently being used in decision-making research. Humans are portrayed as intellectual cripples, limited in their capacity to think, and biased by cognitive processes that interfere with rational decision making (e.g., Dawes, 1976; Slovic, Fischhoff, & Lichtenstein, 1976). They are oversensitive to variables that are not

[2]It is also true that when holistic stimuli are used (e.g., Einhorn, 1974; Phelps & Shanteau, 1978), the subjects are often asked to evaluate the levels of the relevant cues as well as to reach a final decision. Thus, the experimenter still defines the relevant cue dimensions for the subject. In addition, it is unclear in which direction the causal arrow flows in such studies. The cue evaluations might well be constructed from an anticipatory decision rather than the decision being caused by an evaluation of the cues. Furthermore, the reactivity of having to make cue evaluations of the holistic stimuli might impose a limit on the external validity of these studies.

included in normative theories (e.g., Kahneman & Tversky, 1972) and undersensitive to variables that are (e.g., Kahneman & Tversky, 1973). They become more variable when given more information (e.g., Einhorn, 1971; Hayes, 1964) and increase their confidence in the accuracy of their judgments when they should not (e.g., Kahneman & Tversky, 1973; Slovic & Lichtenstein, 1971).

If we eliminate the derogatory tone of these conclusions, what is left is a simple descriptive statement suggesting that decision makers are sometimes responsive to task characteristics that are not specified by prior normative or theoretical conceptions (Olson, 1976) and that researchers do not know when such oversensitivities will emerge. In some tasks, certain variables have smaller effects than expected; in other tasks, the effects are larger than expected. Put differently, there are no theories to tell us when people will be Bayesian, when they will average, when they will add, when they will be subjective-expected-utility maximizers, when they will be sufficiently sensitive to characteristics of data samples, when they will show appropriate hindsight, when they will retrieve information from memory that is not typical but is actually representative, when they will know what they do not know, and so on. What features of tasks control when and which of these many different processes will have causal effects on decisions? How and when might these different processes interact?[3]

If features of simulated decision tasks that are not included in the existing models of basic processes are controlling the subjects' decisions, even to some extent, then one has at least two options. The first is to broaden current models to include these features of the task and thus manage to retain the assumption that simulated decision tasks tap basic processes. This seems to be the popular response. Invoking heuristics (cf. Pitz, this volume), biases, transformation of variables previously thought not to require them (cf. Kahneman & Tversky, 1979), and postulating several decision strategies where before there was only one (cf. Wallsten, this volume) are the frequently used strategies for explaining results that do not fit an expected outcome.

A more radical alternative is to change one's view of decision making. Rather than think of decision making as controlled by a few basic processes that can be discovered by studying a limited and arbitrarily selected set of decision tasks, one could assume that decision rules and processes are *created* to fit the specifics of each particular decision task. In this view, features of a decision task and of measurement procedures (cf. Fischhoff, Lichtenstein, & Slovic, this volume) that have little or no theoretical relevance to the researcher might be expected to

[3]Another way of speaking about the fact that decision-making processes seem to be highly task specific is to say that the causal relationships between specified cues and measures of decision making vary with the context. It cannot be concluded that base-rate information is ignored because sometimes it is not (Kassin, 1979; Wells & Harvey, 1977). It cannot be concluded that sample size has little or no effect on decisions because sometimes it does (Olson, 1976). In short, causal relationships may be less consistent over minor variations in the nature of decision tasks than is generally believed.

determine, at least in part, the results one observes. After adopting this view, one would not be surprised to find that features of tasks, such as the context, the order in which information is presented, the salience of different cues, the number of times a decision is made, the response scales used, the way in which the task is described, the abstractness of the information, the amount of time given to decide, and so on, might affect the decisions of subjects. Rather than "explain" these effects by assuming the existence of all sorts of cognitive limitations and biases, one might think of people as continually shifting their strategies to meet the demands placed on them by contrived decision tasks.

## Comparison of Laboratory Simulations With Real-World Tasks

Several studies that have compared the results from simulated decision tasks to results obtained from unobtrusive (Webb et al., 1966) observations of the decision situations being simulated have recently emerged. These provide a different and more direct source of evidence against the utility of the view that most decision tasks tap basic decision processes.

*Bail Setting.* In a study of bail setting (Ebbesen & Konečni, 1975, in press a), we presented San Diego County judges, who had had first-hand experience with bail setting, with simulated cases and asked them to set bail, in dollars, *exactly* as they would if the case were a real one. The cues that the judges were to use in reaching their decisions were presented in decomposed form on a sheet of paper. Following a brief description of background information (which included the same charge for all cases), the following information was presented: (1) prior record; (2) the extent to which the accused was tied to the local area (for example, owned a home, was employed, and was married); (3) a dollar amount recommended by the district attorney; and (4) a recommendation by the defense attorney, also in dollars. Prior observation of *actual* bail hearings showed that these cues were typically presented to the judges prior to their decisions and that little other information was presented or otherwise available to the judges. Interviews with the judges and official bail-setting guidelines both suggested that local ties would be the most important factor in the decision. The levels of the various cues were organized so that they formed a complete factorial design. Analysis of variance of the bail amounts indicated that all but the defense attorney's recommendation had significant effects, and in obvious directions. There were no interactions. The local ties variable did indeed account for the most variance, by far.

Taking an untypical next step, we also trained observers to code, *unobtrusively,* the levels of the same variables, as well as to record the final amount of bail set in *actual* bail hearings presided over by the *same* judges used in the simulations. The judges were completely unaware that these observations were

being made. The reliability of the coding was virtually perfect. Multiple regression analyses of these naturalistic data indicated that it was possible to account for almost all of the variance in the bail decisions (95%) with the same four factors manipulated in the simulation (plus the severity of the crime). More importantly, a quite different pattern of results emerged. The district attorney's recommendation accounted for the most variance; the defense attorney's recommendation was significant; local ties accounted for a nonsignificant portion of the variance; several interactions emerged. Two related interpretations for the differences in the results between the simulated and the actual bail decisions are: (1) that the range of values of the various cues was different in the two studies and (2) that the interval scale spacing of the levels of the cues used in the multiple regression did not match the judges' subjective spacings of the cue levels in the simulation. To test both of these possibilities, a dummy variable multiple regression that utilized only those cases in which the cues took on values very close to those used in the simulation was performed. The results indicated that the district attorney's recommendation was able to account for almost all of the predictable variance in this data set.[4]

In short, the picture of the judges' bail-setting strategies that emerged from the simulation was quite consistent with the bail-setting guidelines; local ties seemed to be the most important factor in the decision. In contrast, analysis of the decisions in the actual bail hearings suggested that judges were primarily influenced by the district attorney's recommendation and that local ties played only a minor role and even then in a direction largely opposite to that found in the simulation! It is of interest to note that the district attorney's recommendation was predicted primarily by the severity of the crime and not by local ties.[5]

*Sentencing of Adult Felons.* As part of the same extensive project on legal decision making in which the previously described bail-setting results were obtained (Ebbesen & Konečni, 1976, in press b; Konečni & Ebbesen, in press a, in press b; Konečni, Mulcahy, & Ebbesen, in press), we have examined the factors that control the sentencing of adults convicted of felonies (a crime punishable by a year or more in state prison). In two simulation experiments, volunteer college

[4] It is possible that the results of the simulation would have been more like those in the actual hearings had severity of the crime been varied as well as the other factors. On the other hand, if the results that are obtained in simulations depend so heavily on including all of the "right" factors as variables, how does one determine what all the right factors are without collecting data in the real world?

[5] Another explanation for the differences between the simulation and the actual hearings is that the severity of the crime (or some other variable) might be correlated with the district attorney's recommendation. The first possibility was assessed by examining the *additional* variance accounted for by the district attorney's recommendation, *after* the crime was included as a predictor. The identical pattern of results emerged. The latter possibility could not be assessed directly; however, observation of the actual hearings suggests that such a factor would be difficult to discover. Even if one or more such factors could be discovered, it is important to note that the resulting picture of bail setting would *still* be very different from that obtained from the simulation.

students were used as subjects. They were asked to sentence people convicted of a felony on a scale from 0 to 25 years in prison. Cues were presented in decomposed form but embedded in a longer "case description." Four cues were manipulated in a complete factorial design: severity of the crime (forgery versus burglary versus armed robbery), prior record (none versus two previous felony convictions), social history (broken home and bad family life versus solid middle-class life), and feelings of remorse about the criminal activity (none versus a lot). All aspects of the two experiments were identical except that one employed a between-subjects design and the other a complete within-subjects design. No interactions were found in either design. All four main effects were highly significant in the within-subjects design. All but the social history factor were significant in the between-subjects design. Severity of crime and prior record accounted for the most variance in both designs, but crime accounted for slightly more in the within-subjects design, whereas prior record did so in the between-subjects design. In short, slightly different conclusions might have been reached had only one or the other simulation study been conducted.

We repeated similar simulation studies with superior court judges and probation officers as subjects. The latter write extensive reports detailing the criminal activity, prior record, social background, and previous legal history of the offender. These reports are given to the presiding judge the day before he is to sentence the offender. The reports conclude with a detailed sentence recommendation. A major purpose of these reports is to provide the judge with background information about the felon and about the crime because the sentencing hearing often provides the judge with his first encounter with the defendant. During the actual sentencing hearing, the district attorney and defense attorney briefly argue for more and less (respectively) severe sentences. The probation officer is usually present but rarely speaks.

Five factors were varied in both experiments. For the judges they were: severity of crime, prior record, method of guilt determination (plea versus trial), social history, and the probation officer's recommendation. For the probation officers, degree of remorse replaced the probation officer recommendation factor. Both experiments employed within-subjects factorial designs. Unlike the college student studies, however, a time-in-prison scale was not used as the dependent variable.[6] Instead, the judges were asked to write down the exact

[6]Contrary to the views of many college students, in most state sentencing systems, the judge does *not* set the number of years in prison. More often than not, the law defines a minimum and/or a maximum sentence. Furthermore, the *actual* length of time that a felon spends in prison is usually controlled by a parole board rather than the judge (Carroll & Payne, 1976; Maslach & Garber, in press; Wilkins, Gottfredson, Robinson, & Sadowsky, 1973). The judge's decision, therefore, is not time in prison but whether to send the felon to state prison or not; and if not, whether the felon should be confined for a brief time (less than a year) to the sheriff's custody (the county jail facilities), be merely released on probation, or be confined to the sheriff's custody and then be released on probation. Had we tried to formulate the decision task with these options for college students, they would not have known what we were talking about. Had we asked the judges to rate years in prison, they would have laughed us out of the chambers.

sentence, with all of the details, that they would give this offender were the description a real case. The sentencing options available to superior court judges are to send the offender to state prison (where he/she remains until released on parole), to confine the offender in county jail (sheriff's custody) for not more than one year and then to follow the jail term with a period of probation (a period of time during which the offender's behavior is restricted and supervised in lieu of confinement), or to merely impose a period of probationary supervision (not more than 5 years per conviction) with no confinement. Other options are available but are rarely used and generally only in special circumstances (Konečni, Mulcahy, & Ebbesen, in press). The probation officer has the identical array of recommendation options available. The results of an analysis of variance of the number of prison sentences given and recommended are presented in Table 2.1. As can be seen, somewhat different patterns of results emerged for the two types of sentencing experts. Although crime and prior record produced the largest $F$ values in both cases, the order of the effect sizes was different. In addition, social history had a significant effect on the probation officer's recommendation but not on the judge's decision, whereas method of guilt determination had a significant effect on the judge's sentencing decision but not on the probation officer's recommendation. There was also a marginally significant crime by prior record interaction for the probation officers.

Comparing these results to the data for college students from the within-subjects design, we find that the students behaved in a manner similar, but not identical, to the judges. Both responded slightly more to the crime than to prior record; however, social history was a significant factor for students and not for the judges. Although the differences between students, probation officers, and judges might be due to any number of factors, sentencing decisions do not appear to be driven by identical rules with identical parameter values in the three instances. Nevertheless, it is possible that in the case of sentencing, the data for the experts are representative of their decision-making strategies in the real world.

TABLE 2.1
*F*-values For Analyses of Variance of the Number of Prison Recommendations by Probation Officers and of Prison Decisions by Judges in Simulated Decision Tasks

| *Factors* | *Probation Officers* | *Judges* |
|---|---|---|
| Crime | 13.70 | 27.68 |
| Prior record | 81.81 | 25.84 |
| Social history | 8.33 | 1.73 |
| Plea/Trial | 1.49 | 8.80 |
| Remorse | 8.94 | |
| Probation recommendation | | 6.69 |
| Crime × Priors | 3.84 | |

TABLE 2.2
Relationship Between Probation Officers' Recommendations and Judges' Sentencing Decisions[a] (Number of Cases)

| | *Judges' Decisions* | | |
|---|---|---|---|
| *Probation Officer Recommendation* | *Prison* | *Probation and Sheriff's Custody* | *Probation Only* |
| Prison | 103 | 32 | 5 |
| Probation and sheriff's custody | 15 | 396 | 42 |
| Probation only | 1 | 34 | 142 |

[a] Total agreement = 87.1%.

The number of factors that can potentially be considered by a judge in actual sentencing hearings is enormous. An attempt was made to code most of these by content analyzing all of the written documents available to the judge prior to the sentencing decision and by recording the stream of verbal interchanges in the hearing (using a time-sampling system in which the identity of the person speaking and the content of the speech were recorded every 10 seconds). Nonverbal factors, such as the appearance and demeanor of the offender, were also recorded. Data for over 3000 cases have been collected. The present results are based upon the 800 or so cases that have been analyzed thus far. A complete description of the methods and coding systems are available in Ebbesen and Konečni (in press b) and Konečni and Ebbesen (in press b).

Of all of the many factors coded, only a very small number accounted for a substantial portion of the variation in the sentencing decisions of judges. By far the best predictor of the sentence was the probation officer's recommendation. Table 2.2 is a contingency table showing the number of cases in which the probation officer recommended prison, probation plus some time in the sheriff's custody, or probation with no period of confinement, and in which the judge gave one of these same three major categories of sentence. As can be seen, the recommendation and the final sentence were in the same category in over 85% of the cases. It is of some interest to note that when there was a discrepancy, judges were slightly more likely to disagree on the lenient (10%) than the severe (6%) side.

When considered separately, we also found that the likelihood of more severe sentences (those involving incarceration) increased as the severity of the crime that the offender had been convicted of increased, and as the prior record of the offender increased. Tables 2.3 and 2.4 show these relationships for broad crime categories and for the number of previous felony convictions.

Another factor that, to our surprise, was also highly associated with the sentence was the manner in which the accused spent the time between arrest and

TABLE 2.3
Percent of Sentences in Prison, Probation and Sheriff's Custody, and Probation Only Categories as a Function of Severity of Crime[a]

| *Crime Category* | *Number of Cases* | *Sentence* | | |
|---|---|---|---|---|
| | | *Prison* | *Probation and Sheriff's Custody* | *Probation Only* |
| Possession of drugs | (106) | 9 | 61 | 30 |
| Forgery | ( 97) | 18 | 47 | 35 |
| Theft | (220) | 14 | 65 | 22 |
| Burglary | (225) | 12 | 67 | 20 |
| Sale of drugs | ( 57) | 14 | 56 | 30 |
| Robbery | (106) | 29 | 62 | 8 |
| Rape | ( 15) | 27 | 67 | 7 |
| Armed robbery | ( 26) | 46 | 54 | 0 |
| Homicide | ( 21) | 62 | 29 | 10 |

[a] Ordering of crimes is based upon average ratings of severity by the same judges whose decisions were observed in sentencing hearings.

conviction—in legal jargon, the status of the offender. Was the defendant released on his or her own recognizance, released on bail, or not released (i.e., remained in jail)? Table 5 presents the relationship between status and the final sentence. As can be seen, being in jail is associated with a greater percentage of severe sentences than being released on one's own recognizance.

Although a few other factors accounted for a small but significant portion of the variance in sentencing, the four factors previously described here were, by far, the best predictors of the sentence. Little is lost, therefore, by ignoring these other factors in the current description.

TABLE 2.4
Percent of Sentences in Prison, Probation and Sheriff's Custody, and Probation Only Categories as a Function of Number of Prior Felony Convictions

| *Number of Prior Felony Convictions* | *Sentence* | | |
|---|---|---|---|
| | *Prison* | *Probation and Sheriff's Custody* | *Probation Only* |
| None | 12 | 60 | 28 |
| One | 10 | 59 | 29 |
| Two | 13 | 62 | 25 |
| Three | 19 | 62 | 19 |
| Four | 30 | 57 | 13 |
| Five and up | 29 | 62 | 9 |

TABLE 2.5
Percent of Sentences in Prison, Probation and Sheriff's Custody, and Probation Only Categories as a Function of the Defendant's Status Between Arrest and Final Sentencing

| | *Sentence* | | |
|---|---|---|---|
| *Status of Defendant* | *Prison* | *Probation and Sheriff's Custody* | *Probation Only* |
| Released on own recognizance | 6 | 61 | 33 |
| Released on bail | 21 | 43 | 36 |
| In jail | 37 | 52 | 11 |

A number of causal explanations can, of course, be generated for the results presented thus far. All four predictors might be differentially correlated with some unmeasured factor that is the single real causal variable. Alternatively, the four factors might be correlated with several different causal factors, each to a varying degree. Although these explanations cannot be discounted, it is difficult to imagine what these other causal factors might be, given the number of variables examined in our work. Still another view, consistent with the simulation work, is that these four factors are cues in the judge's decision and are, therefore, all causally important.

A somewhat different view of the process is to assume that the variables are related to each other in a causal chain (Heise, 1975). Thus, it might be that only one or two of the four factors are direct causes of the sentence and that other factors are causes of these causes. Several temporal features of the system make certain chains less likely than others. For example, it is always the case that prior record, status, and severity of the charge at conviction are determined earlier in time than the probation officer's recommendation and the judge's sentence. Although it is not impossible to imagine a view of the system in which the final sentence caused prior record (say, via selective reporting or alteration of rap sheets on the part of probation officers), the occurrence of activities such as these was very unlikely in the studied circumstances. Accepting the temporal order, for the moment at least, as useful causal evidence, it is possible to construct several reasonable causal models relating the five variables to one another. Fig. 2.1 presents a diagramatic representation of three such models. In the top model, prior record, severity of crime, and status are assumed to be direct causes of the probation officer's recommendation. But, these variables are assumed to have no direct causal link to the sentence decision. Only the probation officer's recommendation is given this distinction.

The second model proposes that the three early factors have direct effects on the probation officer's recommendation and on the judge's decision, but the latter

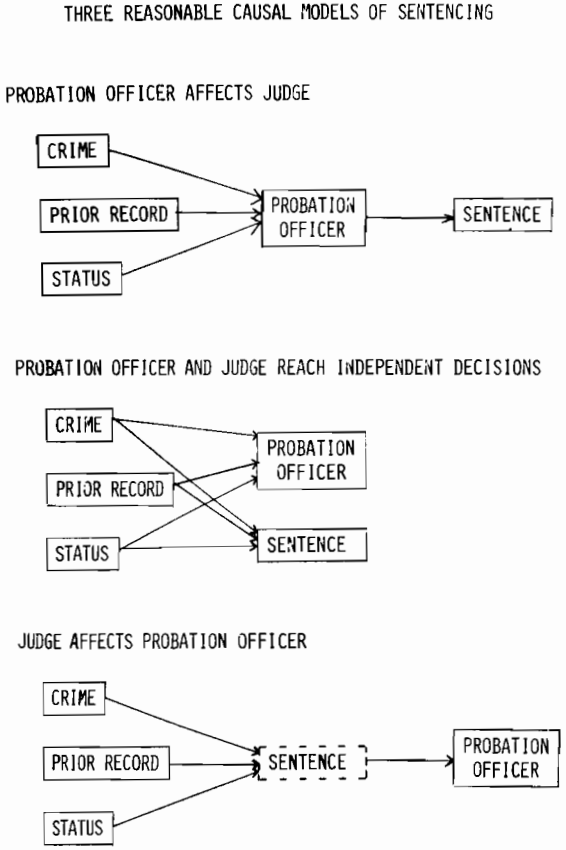


**FIG. 2.1.** Three causal models for the relationships between crime, prior record, status, probation officer recommendation, and sentence. The arrows represent the existence and direction of causal relationships between a pair of variables.

are not causally related. In this view, the high agreement shown in Table 2.2 between the probation officer and the judge is assumed to be a spurious consequence of the fact that both variables are being caused in the same manner by the same set of prior variables.

The third model actually reverses the temporal order of events and argues that the probation officer's recommendation is directly caused by the judge's sentence, which is, in turn, caused by the three prior factors. One reasonable interpretation of this temporal reversal is to assume that the judge is committed to a specific sentence agreement made between the district and defense attorneys in exchange for a plea of guilty (contrary to the popular opinion represented by Perry Mason television shows, over 85% of all felony convictions are obtained as a result of guilty pleas, rather than as a result of jury trials) and that the probation

officer writes his/her report and recommendation in correct anticipation of the judge's decision and is motivated to match and justify this sentence agreement.[7]

Each of the above causal models implies that observed cell frequencies in the five-way data table (crime by prior record by status by probation officer recommendation by actual sentence) should be due to a particular set of "main" and "interaction" effects. These effects can be represented as log odds ratios in a linear model. Specifically, each variable and interaction between variables adds to (or subtracts from) the odds (in logarithms) of an observation falling into a given cell. The causal relationships (arrows in Fig. 2.1) that are assumed by a particular causal model constrain the set of odds ratios (effects) that are to be used to predict the observed frequencies. For example, the first model in Fig. 2.1 assumes that the two-way "interaction" between the three prior variables and actual sentence are all zero (there are no arrows between these variables), whereas the second and third models do not impose this constraint. The predicted (by a given causal or log-linear model) cell frequencies can be estimated from appropriate marginals and then tested against the observed frequencies by an ordinary $\chi^2$ test of fit. Because the details of this analysis approach have been described elsewhere (Ebbesen & Konečni, in press b; Goodman, 1973), we simply report the major results here. When each of the three models outlined in Fig. 2.1 were fitted to the observed frequencies, the first model was best able to account for the data. In addition, the fit of this model to the observed pattern of frequencies was excellent, even in absolute terms ($\chi^2$ likelihood ratio (189) = 136.44, $p > .5$).

Further evidence in support of the first model was obtained by examining the relationship between the prior variables and the sentence when the covariation in the probation officer's recommendation was partialled out and by examining the relationship between the prior variables and the probation officer's recommendation when the covariation in the judge's sentencing decision was partialled out. In the latter analysis, prior record was still significantly associated with the recommendation. In the former analysis, however, all relationships with the sentence were nonsignificant.[8]

Because sentence agreements are not obtained in all cases, it was possible to examine the relationship between the probation officer's recommendation and

---

[7]It should be noted that given the method used to assess the utility of these three models, the third model is isomorphic with the causal model assumed by the previously described factorial simulations, namely, that severity, prior record, status, and the probation officer's recommendation are all direct causes of the judge's final decision. One need merely reverse the direction of the arrow between sentence and probation officer in the third model in Fig. 2.1 to see that this is so.

[8]This test assesses whether the variation in the three prior variables can "get through" to the last factor when the intervening variable is held constant, statistically. If the first model is correct, holding the probation officers' recommendations constant should prevent crime, prior record, and status from being related to the sentence. This is exactly what happened. The logic of the reverse test is similar and supported the present view.

the sentence for just those cases in which no agreement was made. If the high agreement between the two decision makers seen in all of the data (Table 2.2) is due to a presentence arrangement with the district attorney, then one might expect the agreement between the judge and the probation officer to be considerably less in this selected sample of cases because the probation officers would not feel constrained to match their recommendations to a preexisting agreement. Table 2.6 shows the relationship for cases in which a presentence agreement was not reached. As can be seen, an equally strong relationship between the recommendation and final sentence was found here as was found in all of the data.

Once again, the picture emerging from an unobtrusive analysis of real-world decisions is different from that obtained from experimental simulations. Although there are other causal interpretations of the naturalistic data previously presented, they are also quite different from the conclusions that would have been reached had we stopped with the simulations. It appears that in the real world the judges respond primarily to the probation officers' recommendations and that case factors have their effects on the final outcome only indirectly by affecting the probation officers' recommendations.

Selecting cases to match the levels used in the simulations supports the claim that decisions in the simulations were based on different rules than those in the real world (Konečni & Ebbesen, in press b). Specifically, (1) the simulations yielded many more sentences involving prison and jail terms than actually occurred in the real world; (2) factors significant in the simulations were not significant (even when considered singly) in the actual hearing (for example, the method of guilt determination was not significant in actual sentences whereas it was in the judge's simulated decisions and remorse was not associated with the actual recommendations of probation officers whereas it did have an effect on their simulated decisions); (3) the agreement between the probation officer's recommendations and the judge's sentencing decisions was much more in the

TABLE 2.6
Relationship Between Probation Officers' Recommendations and Judges' Sentences When no Preplea Bargain Agreement was Made[a]

| | *Judges' Decisions* | | |
|---|---|---|---|
| Probation Officers' Recommendations | Prison | Probation and Sheriff's Custody | Probation Only |
| *Prison* | 18.8 | 5.3 | .4 |
| *Probation and sheriff's custody* | 1.2 | 45.7 | 5.7 |
| *Probation only* | 0.0 | 3.3 | 19.6 |

[a] N = 245. Numbers in table are percent of total cases falling in each cell. Total probation officer-judge agreement = 84.1%.

real world (over 85%) than in the simulations (approximately 40%); and (4) the relative importance (variation explained) of the several factors were different (the probation officer's recommendation was the most important factor in the actual decisions, whereas it was one of the lesser factors in the simulation). Another important difference that emerged is that the simulation results implied that the judges and the probation officers were responding to somewhat different cues (see Table 2.1). However, when the real-world results were analyzed, treating the judge and probation officer as independent decision makers, we found that their decisions seemed to be responsive to virtually identical cues and in the same order of importance, not surprisingly, given the high agreement between them. Still another difference, not yet discussed, is that several interactions between cues were detected in the real-world data (for example, crime by prior record, crime by status) that did not emerge in the judges' simulation (although, in a rather different pattern, a marginally significant crime by prior record interaction was found for probation officers in their simulated decisions). Finally, the best-fitting decision rules for the judges were different in the two studies: In the real world, judges seemed to decide simply on the basis of the probation officers' recommendations, whereas in the simulation, they seemed to linearly combine crime, prior record, method of guilt determination, and the probation officers' recommendations.

*Automobile Driver Behavior.* A study of driver decision making has suggested that experimental simulations can yield different results than those obtained from unobtrusive observation in decision situations involving risk (Ebbesen, Parker, & Konečni, 1977). In this study, we found that drivers seemed to decide whether to turn in front of an oncoming car (or let it go by before turning) at a T-intersection on the basis of the *temporal* gap between the driver's car and the oncoming car. When we attempted to construct a "holistic" simulation of this situation in the laboratory, we found that experienced drivers seemed to respond, separately, to the speed and the physical distance of the oncoming car rather than to a direct perception of the temporal gap. Had we only conducted the laboratory simulation, we would have concluded that distance and speed were being independently evaluated and weighted, and then configurally combined to reach a decision. Instead, the field data suggested that the turning decision was a direct and simple function of the temporal gap between the two cars and that the drivers were merely applying a simple threshold rule to the temporal gap dimension in deciding whether to turn.

*Judging Swine.* Phelps and Shanteau (1978) have reported that livestock judges took many more cues into account in their judgments of swine when the cues were presented as a fully crossed factorial design in decomposed form than when pictures of swine, rather than feature lists, were evaluated. Although a major conclusion of this work (with which we agree) was that the currently

popular conception of decision makers as being limited in their capacity to take a large number of factors into account in making decisions is in error, the fact remains that differences in decision results and, therefore, in the apparent underlying process were obtained across the two tasks. One among many reasonable alternative explanations for these differences is that the factors were correlated in one task and not in the other. As we (Ebbesen, Parker, & Konečni, 1977) and others (Brunswik, 1956) have argued, whether the correlations between potential cues deviate from zero may well be yet another feature of decision tasks that alters the decision strategy people use.

## Arguments For and Against the Task-Specificity Approach

A number of arguments might be raised against the evidence cited in the previous section. It is conceivable that the simulations were poor representations of the decision tasks being simulated and that had they been better, the results from the different procedures would have been more similar. Although this argument cannot be refuted until the "better" simulations are conducted, two comments about it should be noted. First, had the real-world data never been collected, no one would have known how "bad" the simulations were. When we first began our present line of research, the real-world data were included as an afterthought. We did not think of the simulations as simulations. Instead, they were designed to provide the *real* causal evidence for what we expected to observe (only as correlations) in the real-world settings. Of course, it could be argued that we, and we alone, are poor at designing simulations. On the other hand, pie diagrams, brief verbal sketches, or a single sentence describing the percentage of people who fall into a certain category do not seem far removed from the simulations that we constructed.

Second, the argument applies equally well, in reverse, to *all* simulation studies. Because data for real-world decision tasks are usually missing from reports utilizing simulation methods, the possibility exists that many of these simulations are also poor representations of the decision tasks they are simulating. Being cautious scientists, the reasonable view is to assume that the results are *not* representative until shown otherwise.

Another argument against the view that simulations generally create task-specific decision strategies is that the real-world decision data we have reported are all correlational and, therefore, solid evidence about real *causal* decision processes can never be obtained from them (Phelps & Shanteau, 1978). Thus, it is possible that the discrepancies are due to our inability to tease apart real from spurious causal relationships in the real-world data. Although this argument has merit in most contexts, we feel that in the present case it lacks strength. It can reasonably be maintained that *all* decision models, whether based on data from

simulations or from observations of real-world events are, in fact, only *paramorphic representations* (Hoffman, 1960) of the *actual* decision processes of the subject. Our models merely simulate—that is, are correlated with—the input-output relationships that we observe (Payne, Braunstein, & Carroll, in press). Even when the claim that deep decision processes are being discovered is buttressed with reaction time, eye movement (e.g., Russo & Rosen, 1975), and/or verbal protocol (e.g., Carroll & Payne, 1976) data, input-output relationships are still being dealt with. One simply has more types of output to consider. After all, people can think about things they are not looking at and speak about things that they would not otherwise think about.

The attack on correlational data is weak for another reason. True experiments do not eliminate the possibility that causal relationships other than those proposed as explanations might be producing the results. The fact that randomization generally breaks the correlation between one variable and all prior variables has absolutely no implications for the correlations between that one variable and all following variables. A given manipulation might create quite a number of mediating variables and processes each of which might play a causal role in the final decision (Costner, 1971). Because these mediating processes might well be correlated with each other, we wind up in a similar position to the researcher dealing with real-world data. The best we can hope for is that our models will describe and *predict* patterns in data.

Whether or not our counter argument is accepted, a review of decision-making research suggests that the specific decision strategies used by subjects are very sensitive to a wide range of task variables. It is possible to argue from this evidence alone that decision tasks do not tap a few simple and basic processes.

It might be argued that there is actually no problem with the results from the laboratory simulations, per se. What needs to be done is to change the way that simulations are thought about. Rather than naively assume that subjects in experiments characterize the decision problems that are given to them in a manner identical to the characterization that our theories and models assume, one should, instead, attempt to discover what the subjects are *trying to do* in the task (Simon, 1969; cf. Pitz, this volume). Having done so, it might be found that the subjects are not playing by the ground rules required by current theories. Furthermore, if tasks were constructed so that subjects perceived them in a manner consistent with theoretical assumptions, simulation results might then provide a much better match to the real-world data. On the other hand, how do we discover what the subjects are *really* trying to do? If the concept of trying-to-do-something is central, then why not assume that it plays an important role in real-world decisions as well? Asking the judges what they were trying to do in sentencing yielded quite an array of responses, even from the same judge. Answers focused on such topics as rehabilitation, recidivism, protection of the public, retribution, deterrence, the extent of guilt, the likelihood of future employment, mental

illness, cost to taxpayers, what was best for everyone concerned, the strength of the evidence, taking everything that is important into account, and so on. Which of these many possibilities represents what the judges are, in fact, trying to do?

Another attack on our position is that we are preaching scientific nihilism. After all, if laboratory tasks create specific, rather than tap into basic, decision processes, then why not assume that real-world tasks also create just as task-specific decision strategies? We would agree with the latter point, but disagree that nihilism is the consequence. What we are suggesting is that in the area of decision-making, the really important truths are to be found in the real world rather than in laboratory simulations, no matter how high the *face* validity of the latter might be. We would prefer to base our conjectures about how people make various types of decisions on observations of *those* people making *those* decisions. We are not arguing that laboratory simulations should be abandoned altogether. There are conditions in which they might serve as useful tools in teasing apart certain questions about the real-world process. Rather than *assume* that the simulations are good, however, one ought to collect sufficient evidence to test whether the constructed tasks have captured the necessary detail of the real world to be *real simulations*. One ought to be required to show that the simulations can mimic data from various aspects of the real world before claiming that one is tapping basic processes.

We are not arguing against the continued search for basic, highly generalized rules of decision making. We are claiming, however, (1) that such rules or models are going to be very hard to discover; (2) that findings from a few laboratory simulations do not establish the generality of a model or process of decision making; and (3), most importantly, that it will be impossible to utilize such basic rules to predict decisions in real-world tasks unless a great deal is known about the task and the decision maker prior to application of the rules, i.e., unless real-world data have already been collected.

The latter claim is a consequence of the conclusion that causal relationships between cues and measures of decision making are not universal but vary over tasks (not to mention subjects). The existence of interactions with task features means that any rule, heuristic, process model, and so on will necessarily have to include parameters whose values are set according to specific features of the decision task. Furthermore, the number of such parameters will almost certainly be very large. Unless real-world data have been collected in an attempt to tie down these unspecified parameters, *a priori* prediction of real-world decisions will necessarily be quite limited.

When one is given a real-world decision task with all of its naturally occurring complexity, the theory must be made to fit the task rather than vice versa. Theoretically irrelevant features cannot be eliminated from the processes by which people make decisions by constructing a task in which such features are held constant.

## SOME GENERAL IMPLICATIONS OF TASK SPECIFICITY

Experimental and methodological procedures for assessing the external validity of causal hypotheses have been described by others (e.g., Brunswik, 1956; Campbell & Stanley, 1963; Rosenthal & Rosnow, 1969; Slovic et al., 1977; Webb et al., 1966). We are by no means the first to raise the question of external validity. Psychologists have been grappling with the issue for years. It simply seemed that the time was ripe to mention the issue once again and in the current context. In part, this is because much of decision-making research seems to have a relatively obvious applied orientation, and yet, little concern about external validity issues has been expressed in the recent literature. We also felt that an emphasis on real-world decisions might focus attention on some neglected issues.

### Appropriate Uses of Laboratory Decision Tasks

If it is agreed that laboratory simulations may not, in general, simulate what they are thought to simulate, then it is reasonable to ask whether there are uses to which laboratory decision tasks might be put other than simulation. One reasonable possibility is to use laboratory tasks not to discover what people do (in general), but rather to arrange demonstrations of what people can or might do (even if only in very restricted circumstances). Thus, it might be of interest to know that a task can be constructed in which decision makers do not respond to base-rate information or in which they are overconfident in the accuracy of their predictions. Although such research seems quite reasonable, an ever present danger is that it can be mistaken for a simulation and that its results will therefore be overgeneralized. For example, several recent studies have shown that people sometimes do use base-rate information (see Kassin, 1979 for a review) and that the use to which sample information is put depends on specific task features (Olson, 1976).

Another use of laboratory decision tasks might be to study the cognitive limitations of decision makers. How many factors can a decision maker take into account? How fast can decisions be made and still be accurate? How much better can experts be than nonexperts? At first thought, such questions seem well suited to analysis with laboratory decision tasks; however, it is quite possible that people's limitations change across tasks. For example, it is generally well known that the number of words that a person can remember from a list varies with the strategy the person uses to remember those words. Similarly, the speed with which decisions can be made depends on the specific nature of the question being asked (Ebbesen & Allen, 1979). Phelps and Shanteau (1978) found that many more factors were taken into account when the factors were uncorrelated than when they were intercorrelated. In short, cognitive limitations may be as specific as decision strategies.

Another critique of the use of laboratory decision tasks to assess cognitive limitations is that a theory of the initial conditions necessary to ensure that people are performing at their limits is presently unavailable. Is 20 dollars a large enough incentive or would the threat of torture push people to greater limits? Should distracting noises be masked with white noise or blocked out entirely with the use of a soundproof chamber? Until agreed-upon answers to questions such as these are obtained, the possibility that current limits might be exceeded with minor task modifications will always be present and generalization to real-world tasks will be difficult.

From a somewhat different perspective, various cognitive limitations and biases that have been demonstrated in laboratory decision tasks often seem to be avoided in real-world settings. For example, the effect that the availability of relevant instances in memory can have on judgments (Kahneman & Tversky, 1973) can be minimized, or eliminated completely, by features of naturalistic decision situations. More specifically, in the case of sentencing, judicial predictions about the future criminal behavior of defendants are not likely to be biased by the judge's tendency to base such predictions on an available subset of the defendants' prior convictions. A "rap sheet" that details the full extent of the defendant's prior record is in front of the judge when he/she sentences the defendant. Thus, memory for the previous criminal behavior of the defendant is externalized. This does not mean that availability plays no role in judicial decisions, however. The availability of relevant facts will almost certainly influence judicial decisions. It is merely that the availability of these facts is determined by other members of the social network who compile those facts rather than by a cognitive heuristic. In short, some real-world tasks may reduce or eliminate the influence of potential cognitive biases. Until appropriate naturalistic data is obtained, it is difficult to know the relevance that various biases and limitations have for real-world decisions.

## The Role of Norms

Another issue that the task-specificity argument raises concerns the use of normative models in decision-making research. As we suggested earlier, many cognitive processes (biases) have been invented recently to explain why people do not behave in accord with predictions from normative models. The evidence for these processes comes largely from studies constructed to be reasonable representations of decision problems to which the normative model might be usefully applied. It is then implied, although not directly, that decision makers are likely to be biased by these processes whenever and wherever they make decisions.

It should be obvious that the very notion of a biasing process only makes sense in the context of a normative model. A given decision outcome cannot be biased unless there is a better one against which to compare it. Because bias is a

comparative concept, the degree and type of bias that one observes necessarily depends on *both* the choice of norm and the choice of observed outcome. If this line of argument is accepted, it suggests the possibility that the biasing processes that have been *constructed* by experimenters and then used to impugn the intellectual ability of people may be specific to the norms to which outcomes are being compared, as well as to the tasks in which the normative violations have been observed.

This dependence on choice of norm would not be a problem if everyone agreed as to what the normative rules should be. Unfortunately, in the real world, consensus is hard to come by. Consider, for example, the fact that the Christian Bible, the Koran, and the Talmud serve as normative devices for a large number of real-world decisions. Are we to convince a nun, about to take her vows of chastity and poverty, that a Bayesian model of the likelihood of life after death might be preferable to the normative view she currently accepts? Or, more realistically, consider the sentencing decision of judges. What should a normative model of their decisions look like? To apply a Bayesian model, for example, requires that we think of the sentence as a prediction of some future outcome. But what should that outcome be? As we noted earlier, interviews with judges suggest that quite a large variety of outcomes might be used: rehabilitation; recidivism; deterrence; perceptions by the offender, by the victim, or by one's colleagues of the sufficiency of punishment; protection of society; agreement with colleagues; feelings of satisfaction on the part of the judge; being appointed or elected to higher office; the response of the media, and so on. To compound the problem, judges simply do not agree on what the appropriate outcomes should be. To make matters even worse, judges typically deny the utility of normative models that do not take into account the fact (as they see it) that every case is different.

Given that all of the aforementioned outcomes are not perfectly correlated with each other, and that different features of cases are likely to predict different outcomes (for example, prior record probably predicts recidivism but not deterrence), the extent to which judges will appear to be biased by sundry cognitive processes is likely to depend on which outcomes the *experimenter* uses in the normative model.

Even in the unlikely event that an agreed upon outcome could be found, there may still be disagreement about the decision rule. Should the likelihood of that outcome be maximized, would a minimum likelihood be sufficient, should the sufficient likelihood vary with the nature of the offense or some other variable, or should the likelihood of the outcome be maximized while trying to keep the likelihood of other outcomes above (or below) specified limits? Thus, the experimenter's choice of rule (as well as outcome) can make an otherwise "rational" decision seem "biased." Cognitive biases may be as much in the mind of the experimenter as in that of the subject.

## Causal Chains in Decision Making

When the real world serves as a source of data, one's view of the typical decision-making process is considerably different from the view that seems common in laboratory decision tasks. In many real-world situations, decisions are actually a part of a larger social system in which the decisions of various people are interrelated in complex ways. When such is the case, it is possible to focus attention on the entire system rather than on one class of participants. The input-output relationships of the system can then be explained by the operation of underlying processes; however, in this instance, the underlying processes are the observable actions of key decision makers in the system and not the unobservable activity of retrieval processes, encoding mechanisms, or decision strategies located somewhere under the skin of the decision maker.

Our own research on decision making in the legal system takes this broader view. Several interesting discoveries emerged because of it. For example, one of the major predictors of the final sentence was the "extra legal" factor: status of the defendant between the time of arrest and the final sentence hearing (see Table 2.4). It appeared that status had its effects by controlling the sentence recommendation of the probation officer which in turn controlled the judge's final decision. But, recall that in our study of bail setting, the amount of bail that a defendant had to pay was controlled by the district attorney's dollar amount recommendation. Because defendants are less likely to be able to afford the usual 10% bail-bondsman fee as the amount of bail increases, and because people who cannot pay bail remain in jail, it is conceivable that the district attorney's bail recommendation, made 2 or 3 days after an arrest, is having a causal effect on the final sentence: a decision being made practically a year after the bail hearing! In short, the decisions of people embedded in a complex social system may be interrelated in ways than can only be discovered by examining the real-world system, *in vivo*.

## SUMMARY

There is considerable evidence to suggest that the external validity of decision-making research that relies on laboratory simulations of real-world decision problems is low. Seemingly insignificant features of the decision task and measures cause people to alter their decision strategies. The context in which the decision problem is presented, the salience of alternatives, the number of cues, the concreteness of the information, the order of presentation, the similarity of cue to alternative, the nature of the decomposition, the form of the measures, and so on, seem to affect the decisions that subjects make. In addition, comparisons of results from simulated and real-world tasks suggest that decision strategies may be task-specific rather than caused by a few basic processes. One conse-

quence of this view is that researchers should provide external validity evidence for claims that causal models derived from laboratory data apply to decisions in real-world settings. The accumulation of such evidence can only serve to broaden our understanding of decision making.

## ACKNOWLEDGMENTS

Preparation of this paper was facilitated by National Institute of Mental Health Grant MH 26069 to Ebbe B. Ebbesen and National Science Foundation Grant GS 42802 to Vladimir J. Konečni.

## REFERENCES

Anderson, N. H. Information integration theory: A brief survey. In D. H. Krantz, R. C. Atkinson, R. D. Luce, P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2). San Francisco: Freeman, 1974.

Anderson, N. H., & Shanteau, J. Information integration in risky decision making. *Journal of Experimental Psychology*, 1970, *84*, 441-451.

Brunswik, E. *Perception and the representative design of psychological experiments* (2nd Ed.). Berkeley, Calif.: University of California Press, 1956.

Campbell, D. T., & Stanley, J. C. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally, 1963.

Carroll, J. S., & Payne, J. W. The psychology of the parole decision process: A joint application of attribution theory and information processing psychology. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976.

Costner, H. L. Utilizing causal models to discover flaws in experiments. *Sociometry*, 1971, *34*, 398-410.

Dawes, R. M. Shallow psychology. In J. S. Carroll and J. W. Payne (Eds.), *Cognition and social behavior*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976.

Ebbesen, E. B., & Allen, R. B. Cognitive process in implicit personality trait inferences. *Journal of Personality and Social Psychology*, 1979, *37*, 471-488.

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

Ebbesen, E. B., & Konečni, V. J. *Fairness in sentencing: Severity of crime and judicial decision making*. Paper presented in a symposium entitled "Seriousness of crime and severity of punishment" held at the 84th Annual Convention of the American Psychological Association, Washington, D.C., September, 1976.

Ebbesen, E. B., & Konečni, V. J. An analysis of the bail system. In V. J. Konečni & E. B. Ebbesen (Eds.), *Social-psychological analysis of legal processes*. San Francisco: W. H. Freeman, in press. (a)

Ebbesen, E. B., & Konečni, V. J. The process of sentencing adult felons: A causal analysis of judicial decisions. To appear in B. D. Sales (Ed.), *Perspectives in law and psychology (Vol. 2): The jury, judicial, and trial processes*. New York: Plenum, in press. (b)

Ebbesen, E. B., & Konečni, V. J. Social psychology and law: Theoretical considerations. In V. J. Konečni & E. B. Ebbesen (Eds.), *Social-psychological analysis of legal processes*. San Francisco: W. H. Freeman, in press (c).

Ebbesen, E. B., Parker, S., & Konečni, V. J. Laboratory and field analyses of decisions involving risk. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, *3*, 576-589.

Ebert, R. J., & Kruse, T. E. Bootstrapping the security analyst. *Journal of Applied Psychology*, 1978, *63*, 110-119.

Einhorn, H. J. The use of nonlinear, noncompensatory models as a function of task and amount of information. *Organizational Behavior and Human Performance*, 1971, *6*, 1-27.

Einhorn, H. J. Expert judgment: Some necessary conditions and an example. *Journal of Applied Psychology*, 1974, *59*, 562-571.

Goodman, L. A. Causal analysis of data from panel studies and other kinds of surveys. *American Journal of Sociology*, 1973, *78*, 1135-1191.

Hayes, J. R. Human data processing limits in decision-making. In E. Bennett (Ed.), *Information systems, science and engineering. Proceedings of the First International Congress on the information systems sciences*. New York: McGraw-Hill, 1964.

Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. Social judgment theory. In M. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes: Formal and mathematical approaches*. New York: Academic Press, 1975.

Heise, D. R. *Causal analysis*. New York: Wiley, 1975.

Hoffman, P. J. The paramorphic representation of clinical judgment. *Psychological Bulletin*, 1960, *57*, 116-131.

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

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

Kahneman, D., & Tversky, A. Prospect theory: An analysis of decision and risk. *Econometrica*, 1979, *47*, 263-291.

Kassin, S. M. Consensus information, prediction, and causal attribution: A review of the literature and issues. *Journal of Personality and Social Psychology*, 1979, *37*, 1966-1981.

Keeney, R. L. A utility function for the response times of engines and ladders to fires. *Urban Analysis*, 1973, *1*, 209-222.

Keeney, R. L. The art of assessing multiattribute utility functions. *Organizational Behavior and Human Performance*, 1977, *19*, 267-310.

Konečni, V. J., & Ebbesen, E. B. A critical analysis of method and theory in psychological approaches to legal decisions. To appear in B. D. Sales (Ed.), *Perspectives in law and psychology (Vol. 2): The jury, judicial, and trial processes*. New York: Plenum, in press. (a)

Konečni, V. J., & Ebbesen, E. B. Sentencing felons. In V. J. Konečni & E. B. Ebbesen (Eds.), *Social-psychological analysis of legal processes*. San Francisco: W. H. Freeman, in press. (b)

Konečni, V. J., Mulcahy, E. M., & Ebbesen, E. B. Prison or mental hospital: Factors affecting the processing of persons suspected of being "mentally disordered sex offenders." In P. D. Lipsitt & B. D. Sales (Eds.), *New directions in psycholegal research*. New York: Van Nostrand Reinhold, in press.

Lichtenstein, S., & Slovic, P. Response-induced reversals of preference in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology*, 1973, *101*, 16-20.

Maslach, C., & Garber, R. B. Decision making processes in parole hearings. In V. J. Konečni & E. B. Ebbesen (Eds.), *Social-psychological analysis of legal processes*. San Francisco: W. H. Freeman, in press.

Naylor, J. C., & Wherry, R. J., Sr. The use of simulated stimuli and the "JAN" technique to capture and cluster the policies of raters. *Educational and Psychological Measurement*, 1965, *25*, 969-986.

Olson, C. L. Some apparent violations of the representativeness heuristic in human judgment. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 599-608.

Payne, J. W. Relation of perceive risk to preferences among gambles. *Journal of Experimental Psychology: Human Perception and Performance,* 1975, *1,* 86–94.

Payne, J. W., Braunstein, M. L., & Carroll, J. S. Exploring pre-decisional behavior: An alternative approach to decision behavior. *Organizational Behavior and Human Performance,* in press.

Phelps, R. H., & Shanteau, J. Livestock judges: How much information can an expert use? *Organizational Behavior and Human Performance,* 1978, *21,* 209–219.

Rosenthal, R., & Rosnow, R. L. *Artifact in behavioral research.* New York: Academic Press, 1969.

Russo, J. E., & Rosen, L. E. An eye fixation analysis of multialternative choice. *Memory and Cognition.* 1975, *3,* 267–276.

Simon, H. A. *The sciences of the artificial.* Cambridge: MIT Press, 1969.

Slovic, P. Analyzing the expert judge: A descriptive study of a stockbroker's decision process. *Journal of Applied Psychology,* 1969, *53,* 225–263.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Cognitive processes and societal risk taking. In J. S. Carroll and J. W. Payne (Eds.), *Cognition and social behavior.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Behavioral decision theory. *Annual Review of Psychology,* 1977, *28,* 1–39.

Slovic, P., Fleissner, D., & Bauman, W. S. Analyzing the use of information in investment decision making: A methodological proposal. *Journal of Business,* 1972, *45,* 283–301.

Slovic, P., & Lichtenstein, S. The relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Psychology Monograph Supplement,* 1968, *78,* No. 3, Part 2.

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

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally, 1966.

Wells, G. L., & Harvey, J. H. Do people use consensus information in making causal attributions? *Journal of Personality and Social Psychology,* 1977, *35,* 279–293.

Wilkins, L. T., Gottfredson, D. M., Robinson, J. O., & Sadowsky, A. *Information selection and use in parole decision making* (Supplemental Rep. 5). Davis, Calif.: National Council on Crime and Delinquency Research Center, 1973.

Zimbardo, P. G., Ebbesen, E. B., & Maslach, C. *Influencing attitudes and changing behavior.* New York: Addison-Wesley, 1977.