AD-A106 143 RICE UNIV HOUSTON TX DEPT OF PSYCHOLOGY F/G 5/10
CHOOSING AMONG ALTERNATIVES WITH UNCERTAIN OUTCOMES: EFFECT OF --ETC(U)
AUG 81 S P KERKAR, W C HOWELL N00014-78-C-0555
UNCLASSIFIED TR-81-2 NL

END
DTIC

CHOOSING AMONG ALTERNATIVES WITH

UNCERTAIN OUTCOMES: EFFECT OF

PRIOR CUING AND ESTIMATION

REQUIREMENTS

Shanta P. Kerkar and William C. Howell
Rice University

Technical Report #81-2

August 1981

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER -81-2	2. GOVT ACCESSION NO. AD-A106	3. RECIPIENT'S CATALOG NUMBER 143
4. TITLE (and Subtitle) Choosing Among Alternatives with Uncertain Outcomes: Effect of Prior Cuing and Estimation Requirements		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Shanta P. Kerkar and William C. Howell		8. CONTRACT OR GRANT NUMBER(s) N00014-78-C-0555
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology Rice University Houston, TX 77001		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR197-050
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE
		13. NUMBER OF PAGES 32
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

uncertainty judgment, choice behavior, probability estimation, frequency estimation, calibration, task effects, predecision processes

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(see abstract)

411624

Abstract

Previous investigations by the authors have shown that judgments of frequency or probability have a facilitatory influence on subsequent predictive choice performance. The present study examined both these judgment processes more closely in an attempt to understand how they affect choices. Five groups of subjects served as emergency vehicle dispatchers for a hypothetical city. The experimental design involved a 2 x 2 factorial combination of response set and estimation task. More specifically, response set, defined as expectation of a frequency or probability response requirement, was crossed with actual response requirement, either frequency or probability estimation, in a between-groups design. In addition to the four resulting experimental groups, a fifth (control) group had no set or estimation requirement at all. Each subject processed frequentistic events (emergency calls defined by type and location), which were programmed to occur in a random order, over three sessions. Then the experimental groups estimated either the frequency (F) or probability (P) of specified events in a manner congruent (FF or PP) or incongruent (PF or FP) with their initial set. Results showed that congruent conditions did not produce reliably better estimates than incongruent ones, but that both frequency estimation conditions (FF and PF groups) were consistently superior to the probability estimation conditions (PP and FP groups). On a subsequent session all five groups made predictive choices among

designated events on the basis of relative likelihood of occurrence. The overall choice performance of the four experimental groups was better than that of the control group, thereby confirming the beneficial effect of estimation on subsequent decisions. However, the congruent groups (FF and PP) outperformed the incongruent groups (PF and FP). This is contrary to what would have been expected if choices merely reflected the quality of estimations. Some possible determinants of predictive choice behavior are discussed for future practical and theoretical consideration.

INTRODUCTION

The fact that the focus of decision-making research has shifted in recent years from a normative to a descriptive emphasis is well documented (Einhorn & Hogarth, 1981; Slovic, Fischhoff, & Lichtenstein, 1977; Wallsten, 1980). However, the rapidly growing body of knowledge that the descriptive orientation has produced is quite diverse both in empirical content and methodology. Although several promising attempts have been made to formulate broad unifying theories (e.g. Hammond, 1980; Wallsten, 1975; 1980), all are still in the earliest stages of empirical verification.

In spite of the lack of an accepted theoretical framework, one generalization that seems to be emerging from the descriptive literature is that task characteristics have an important bearing on decision strategies. Typical support for this conclusion is afforded by studies in which cue utilization, overt judgment, or decision is shown to be affected by the manner in which cues are formatted or presented to the subject (Einhorn, 1970, 1971; Mertz and Doherty, 1974; Phelps and Shanteau, 1978; Slovic and Lichtenstein, 1971; Slovic and MacPhillamy, 1974).

Given that task structure plays a vital role in decision behavior, then, an obvious first step toward understanding this relationship would be the identification and classification of relevant task components--the taxonomic approach. However, a purely structural taxonomy would have limited usefulness, since it is the interaction between the decision-maker (DM) and the task, not the task itself, that shapes the resulting behavior (Dawes, 1975). The DM probably has a variety of information-processing options or strategies in his repetoire for dealing with most decision problems, but which one he invokes in any particular case may well depend upon the task structure. Therefore, a more promising approach to classification would seem to be one focused on the cognitive implications of task features.

Howell & Burnett (1978) followed this logic in proposing a taxonomy for at least one general aspect of decision problems: uncertainty. Basically, they attempted to distinguish some key task parameters in terms of presumed links with underlying cognitive processes. A foremost objective of this taxonomy was to unravel possible conceptual differences among what are commonly regarded as equivalent measures of uncertainty.

In an earlier study we validated one of the predictions derived from this taxonomy, showing that the impression of uncertainty could be directly related to the response required of the subject (Howell & Kerkar, 1981). The assumption there was that a frequency estimation response, which pertains to actual occurrences of repetitive events, is likely to produce a better "calibration" of uncertainty than is a future-oriented probability response. Our results confirmed this hypothesis. We also tested the extent to which the estimations affected subsequent predictive choice performance and found that subjects who made either probability or frequency estimates made better predictive choices than those in a control group who made neither type of estimation. Moreover, the data suggested that the frequency estimation task may produce more beneficial effects on choice performance than the probability estimation task.

We concluded, therefore, that prior estimation has a facilitatory effect on decisions, although it is not clear exactly how the influence operates. One possibility is that it represents a set effect, with estimation cuing the DM to pay particular attention to the frequentistic/probabilistic property of events during encoding of choice events. Previous literature has shown set to be a powerful factor in a wide range of behaviors including perceptual response (Helson, 1948, 1964), problem solving (Luchins, 1942), and even impression formation (Asch, 1946). A second possible explanation is that rather than controlling the actual <u>encoding</u> of events, estimation produces a convenient "summary code" at the time of retrieval which is then referred to during the choice task. Without it, the DM must resort to whatever "raw data" he has stored when faced with a decision. If frequency estimation yields a more accurate picture of what has transpired than does probability estimation, as the data suggest, then perhaps this bias is carried over to the choice task as well. Support for this interpretation follows the suggestion by Einhorn & Hogarth (1981) that judgment may reduce the uncertainty inherent in choice by requiring prior deliberation and an evaluation of evidence.

Without some indication of the processes involved in estimation, of course, it would be impossible to choose between the cuing and the "summary code" hypotheses: both account nicely for the obtained relationship. The major objective of the present study, therefore, was to examine with greater precision the frequency and probability estimation processes, and to explore the implications of process differences for predictive choice (decision) performance.

Considering first the matter of estimations, it is clear that a necessary condition for the production of a veridical estimate of either frequency or probability is the encoding and storing of past event occurrences. Research in the area of memory has shown that subjects tend to employ encoding strategies consistent with anticipated retrieval requirements: that is, they store information in a form designed to facilitate later retrieval (Tversky, 1973; Tversky & Teiffer, 1976). In the present context, this suggests that the quality of an estimate would depend partly upon whether the subject anticipated a frequency or probability response. Although some studies have failed to show a cuing effect in the estimation of simple word frequencies (Hasher & Zacks, 1979; Howell, 1973), recent work by Marques and Howell (1979, 1980) did find it in a more complex task setting much like the present one. Therefore, response cuing seems a plausible means of inducing differences in the encoding of to-be-remembered information. Cuing a frequency response should accentuate those features of event occurrences that are cognitively most germane to the estimation of past frequency; cuing a probability response should do the same for those associated with the estimation of future probability. If the same processes underlie both, then it should not matter which cue is associated with which response. If they are different, then it should matter a great deal: cuing consistent with the response (e.g. probability cue-probability estimate) should produce better estimates than inconsistent cuing (e.g. probability cue - frequency estimate).

Remember, however, that our previous study suggested an inherent superiority of frequency over probability estimation performance. A convenient way to extricate this difference from the possible cuing effect is simply to cross the two in a factorial design with estimation performance as the dependent variable. Thus identified, the major processes involved in estimation could be related to subsequent predictive choice performance. That is, cuing consistency should control the quality of encoding; response mode should control the choice of information used in forming a "summary code" (and hence the amount of any bias present in it). Should either or both of these factors have a significant effect on estimation performance, the correspondence between that effect and predictive choice performance would indicate the relative contribution of the two processes; and this information, of course, would go a long way toward determining which hypothesis of the estimation-decision relationship is more tenable.

The present study, therefore, was similar to our previous one except that a between-groups cuing manipulation was incorporated into the design. In addition, the predictive choice task was expanded to include both two-alternative and three-alternative problems with the aim of providing further insight into the processes linking estimation and choice. The rationale for the latter manipulation was that if the DM uses the same stored information in choosing that he does in estimation, then it should make no difference whether a particular pair of alternatives is presented separately (2-choice situation) or embedded within a triad (3-choice situation). On the other hand, if the probability estimation mode induces a bias of some sort, then its effect should be amplified in the 3-choice situation.

Method

Task Scenario (acquisition). The basic task used to expose subjects to a complex array of event frequencies was that of emergency vehicle dispatchers for a hypothetical city (see Howell and Kerkar, 1981, for a more complete description). The events were emergency calls of various kinds to each of which the subject was required to give some response before the next event appeared. There were 36 distinct kinds of calls, defined, for purposes of face validity, in terms of type of emergency (fire vs. police), veracity (true vs. false alarm), and location (nine city sectors). Each type of call was programmed to occur at a frequency between 0-10. The assignment of these frequencies to events was randomized (see Table 1), and

---
Table 1 about here
---

the resulting distribution of 75 calls was assigned randomly to positions within a sequence for presentation. A session or "shift" was defined by the occurrence of these 75 events in a particular order. The sequence, but not the distribution, of events was varied randomly over shifts. This made it possible for subjects to gain familiarity with the distribution of calls by type, veracity, and location: the "event generator," as it were, was stationary and therefore learnable.

The task was explained to the subject using a map identifying the location and principal land use (e.g. residential, commercial) of each sector. This literal format was also used for displaying input and response information to the subject as it occurred during each session. The entire task, in fact, was carried out on a TRS-80 microcomputer (see below).

At the start of each shift, the subject was given a fixed number of police and fire vehicles. He was required to distribute these total resources (vehicles) among the sectors in what he considered to be the most advantageous fashion for later allocation to particular emergencies. Then the emergency calls appeared one at a time. On receiving a call, he responded either by dispatching a vehicle immediately or by verifying the call. Immediate feedback was given as to whether the call was a true emergency or a false alarm. The entire distribution of calls was presented in this manner. The responses were scored according to a cost-payoff scheme which had the following general features: (a) a correct response, i.e. dispatching a true emergency or verifying a false alarm, was scored positively, (b) verification became more desirable as the false alarm rate increased, since dispatching a call reduced the number of available vehicles, (c) a penalty was assessed for any responses to calls that occurred in sectors which had no vehicles remaining. The cumulative score was displayed continuously as an indication of the subject's performance during a session. It must be noted that this decision performance was itself of little direct consequence, although the subject was not made aware of this fact. Its primary function was to make the task intrinsically motivating and maintain the subject's attention (as well as to verify these conditions).

Estimation Task. At the end of the three acquisition sessions, the experimental subjects (but not the controls) provided either frequency or probability estimates in accordance with their particular group assignment. Thus, they were queried about the frequencies with which certain events had appeared (frequency estimation), or the chances that the same events would happen in the future (probability estimation).

Predictive Choice Task. All subjects (experimental and control) were required to make decisions on the fourth session. Each subject was again presented with a map of the city, but now either two or three potential calls appeared simultaneously. For each set of alternatives, he had to choose the event that he considered most likely to occur next.

The selection of events for presentation was directly related to the frequency of their occurrence on the acquisition task. It will be recalled that six frequencies (0,1,2,4,8, and 10) were assigned to the distinct event categories, which were defined by type of call, veracity and location (see Table 1). All 20 unique combinations of these frequencies, taken three at a time, were used for simultaneous presentation in the three-choice condition. The decision on which particular events with these frequency characteristics to display in each case was made randomly. The event pairs for the binary-choice condition were obtained directly from these 20 three event combinations. Since each combination yields three pairwise comparisons, 20 X 3 = 60 binary choice pairs were available from this source. Using the frequency combinations so generated, selection of event pairs having these properties was determined from the choice triads. Eighty such pairs were constructed. It should be noted that by generating the pairs out of the triplet alternatives, it was possible to compare specific choices embedded within either the two-or three-choice problem form.

Experimental Design. Response set, defined as the initial expectation of an eventual frequency (F) or probability (P) estimation requirement, was crossed with actual response requirement, also frequency or probability estimation, in a 2 X 2 factorial design. Each of the resulting combinations was the basis for an experimental group. In addition, there was a fifth (control) group which had no estimation responsibilities or set at all (See Table 2 for the complete design). The four experimental groups made estimates congruent (FF or PP) or

----

Table 2 about here

----

incongruent (FP or PF) with their initial set after the third session; all five groups (FF,PP,FP,PF, and C) made two-choice and three-choice predictions during the fourth session. The critical comparisons in this design were: (a) the quality of the estimations obtained under the congruent/incongruent conditions, and (b) the predictive choice performance of all five groups under the binary and three-choice conditions. The accuracy of choices made by the control group provided a base level for comparing the effect of the experimental manipulations. All of these were between-group comparisons except for the two- vs. three-choice format used in prediction.

Subjects. The 45 subjects were recruited from undergraduate psychology courses in partial fulfillment of an experimental requirement. In addition, they were paid a bonus based upon their predictive choice performance which averaged around $3.50 per subject. Assignment of subjects to the five groups was randomized with the restriction that the groups be of equal size (n = 9).

Procedural Details. Each subject served individually on all four sessions which were conducted over two consecutive days. The two sessions per day lasted about 30 minutes each and were separated by a brief rest period during which subjects were allowed to leave the experimental cubicle. All subjects performed the dispatching (acquisition) task for the first three sessions, after which the experimental groups made their frequency or probability estimations. The fourth session was devoted entirely to the predictive choice task for all groups. The temporal sequence of tasks is summarized in Table 3.

Table 3 about here

All subjects were given detailed procedural instructions at the start of the first session, with special emphasis on the importance of their ability to show a steady improvement in their allocation (dispatching) performance. The purpose of this emphasis was to encourage their attending to and processing the incoming calls. The appropriate response set was induced in the experimental subjects by further instructions which forewarned them of a frequency estimation task (FF and FP groups) or a probability estimation task (PF and PP groups). Detailed instructions for both estimation tasks were presented at the end of the third session, and those for the predictive choice task at the beginning of the fourth session.

The subject was seated in an experimental cubicle in front of a TRS-80 microcomputer. The CRT screen provided a continuous display of a map of the city zones, an indication of the available resources, and a record of his cumulative allocation decision score. Each call appeared on the screen in the appropriate location, and the subject entered his response via the computer keyboard. Immediate feedback regarding the outcome of the response was provided on the display.

After the third session, subjects in the four experimental groups were given frequency or probability questionnaires. The items in both of the questionnaires pertained to the occurrence of specific event types and type-location or type-veracity-location combinations. Some illustrations of these items are presented in Table 4.

Table 4 about here

The fourth session was devoted exclusively to the predictive choice task, which was carried out by all groups under identical instructions. The subjects were paid a bonus of 10 cents for every correct prediction, and while they were not given direct feedback on their responses, the cumulative amount of the bonus was displayed occasionally to maintain motivation. They were also required to rate their confidence in each prediction using a 5-category scale which ranged from "randomly guessing" to "absolutely certain." The instructions emphasized that the ratings had no bearing whatsoever on the recorded accuracy of the responses, the bonus earned, or subsequent choices. (The ratings were, in fact, only of peripheral interest in the study.)

After the fourth session, the subjects were paid the bonuses earned, and were debriefed in writing regarding the purpose of the study.

RESULTS AND DISCUSSION

Estimation. Since the event occurrences were generated by a stable process, the subjects' frequency on probability estimates could be compared directly to the actual or "objective" values to assess their accuracy. We used three measures to describe different aspects of the quality of estimation performance: (a) the unsigned error score or deviation of estimates from the absolute reference values (calibration), (b) the % error index which expresses the unsigned deviations as percentage of error relative to the actual-frequency or probability values, and, (c) the correlation of estimates with objective values (discrimination coefficient). It must be noted, however, that only items #4 and #5 provided a sufficient number of judgments to make the computation of the correlation coefficient meaningful. Therefore we relied mainly on the two error scores for overall comparisons. It is also important to note that, although related, these measures yield somewhat different information. The unsigned deviations reflect the accuracy of the absolute value of judgments, whereas % error adjusts this accuracy in proportion to the actual base frequency or probability values. Thus a given magnitude of unsigned error at a small base value (frequency or probability) appears considerably larger than at a large base value for % error. For example, an unsigned error of 1 would be represented as 100% error at frequency 1, but only as 10% error at frequency 10.

Considering the mean error scores across all items (see Table 5), the most noteworthy finding is that performance was obviously controlled by the response required and not by the

---

Table 5 about here

---

prior set. The PF and FF groups performed similarly, as did the PP and FP groups; however the former (frequency response) pair was vastly superior to the latter (probability response) pair on both error measures. This conclusion was supported by analyses of variance results. For the unsigned error index, the main effect of response requirement was significant at F (1,32) = 71.03, $p < .001$, while that for prior set was not,

$F(1,32) = 1.60$, $p < .210$.

For the % error index, the pattern was the same: the response requirement effect was highly significant, $F(1,32) = 121.45$, $p < .001$; the set effect, though closer to significant, failed again at $F(1,32) = 3.08$, $p < .090$. Equally as important as these main effects, the set x response requirement interaction was clearly nonsignificant on both measures: $F(1,32) = 1.42$, $p < .240$ on the unsigned error index; $F(1,32) = 2.75$, $p < .110$ for % error.

It would appear, then, that "setting" subjects for one response mode and then requiring them--unexpectedly--to respond in the other (i.e. the PF and FP groups) produced the same quality of estimations as did the congruent set-response conditions (i.e. the PP and FF groups). This would suggest either that specific response cuing does not induce differential encoding of events as they occur, or that if it does so, the different "records" are equally useful for either type of estimation response. On the other hand, the fact that consistent response mode differences occurred once again (virtually identical to those in Howell & Kerkar, 1981) implies that frequency and probability estimation draw upon different stored information--information that must be encoded under both kinds of set. The most plausible explanation, therefore, is that people maintain similar "frequency records" under either set, but that they rely more heavily on this information relative to other considerations (such as _recent_ or _prominent_ events, or beliefs about event causation) when a frequency response is required. This is not to suggest that the encoding of event frequency information is necessarily _automatic_, or totally impervious to set, as some have argued (Hasher and Zacks, 1979). Marques and Howell (1979, 1980) have shown that quality of frequency estimation is dependent upon attention during the encoding of events. We are simply proposing that both frequency and probability "set" induce a similar level of attention during the event encoding process.

While the superiority of the frequency response mode confirms our earlier results as noted above, it still leaves open the question of exactly how the estimation tasks differ. Our contention that the probability response represents a qualitatively different process is still viable, but the exact nature of that process--particularly the information upon which it draws--is still not clear. Although not specifically directed toward this issue, the present item-probe manipulation does provide a partial answer. It will be recalled that estimations were required for different kinds of items (see Table 4). Thus comparing relative performance on specific kinds of information provides some insight into the underlying processes.

As would be expected, some kinds of probe items were handled better than others by subjects in all groups. The main _items_ effect was significant in both analyses: for unsigned error,

$F(4,128) = 12.25$, $p < .001$; for % error, $F(4,128) = 54.49$, $p < .001$. Since the interaction of this variable with response requirement was also statistically significant--$F(4,128) = 26.89$, $p < .001$ for unsigned error and $F(4,128) = 51.28$, $p < .001$ for % error--the means for the various items must be considered separately for the frequency and probability response groups. This is not necessary, however, for the various set conditions since the set x probe item interaction did not even approach significance: $F(4,128) = .51$, $p < .730$, and, $F(4,128) = .76$, $p < .550$ for unsigned and % error. A summary of the item-related differences is presented in Table 6. Note that the

Table 6 about here

error scores of FF+PF groups and FP+PP groups are combined and described as frequency and probability estimation performance respectively.

The robustness of the superiority of frequency over probability judgments is evident from lower error on both indices on all five items (except % error on #1). It will be recalled that the items pertained to specific event types, event-location and event-location-veracity combinations. Since they probed such diverse information, the significant item effect is to be expected. What is more interesting is the item X estimation task interaction. The pattern of error scores in Table 6 clearly shows that the discrepancy between the performance of both groups is a function of the nature of the probes. Moreover the individual error indices themselves suggest rather distinct trends on estimations, since different aspects of performance are highlighted.

Considering first the unsigned deviations for frequency judgments, the best estimates are obtained on items #4 and #5. However, the ordering of mean error is almost reversed on the % error index with performance on #4 and #5 being the worst. Given that these items involved frequencies of considerably smaller magnitude than #1-3, even slight inaccuracies would be inflated on a relative error index such as % error. In contrast, the moderate amount of unsigned error on items that entailed higher frequencies (#1-3) remained unchanged on % error.

Such an anomaly is not evident in probability judgments. In fact, both the error indices show perfect correspondence: the greater unsigned error on #4 and 5 relative to #1-3 is just amplified in the % error measure.

It appears, then, that the superiority of frequency over probability judgments all but disappears when events are defined very broadly and thus absolute frequencies are relatively large. (e.g. when the subject is asked to estimate total police calls for all locations). The greatest discrepancy occurs on probes related to finer categories where frequencies tend to be smaller (#4, 5). This could well reflect an overconfidence bias in probability judgments that produces a gross overestimation of smaller values (Hogarth, 1975; Lichtenstein, Fischhoff & Phillips, 1977).

The probe-category distinction takes an even greater meaning in light of the correlation data which, it will be recalled, were necessarily limited to the more narrowly defined items (#4 and 5).

Table 7 about here

In Table 7 it is clear that the superiority of the frequency response mode appeared only on probe category #5. This difference was corroborated by a significant probe X response interaction: $F(1,32) = 9.32$, $p < .004$. The overall probe category difference (#4 vs. #5) was also significant, $F(1,32) = 62.96$, $p < .001$, whereas, in contrast to the error analyses, the overall response-requirement effect was not, $F(1,32) = 2.32$, $p < .120$. Considered together, the deviation and correlation results provide additional support for the position that it is a constant <u>bias</u>, not just less <u>precision</u>, that makes probability judgment inferior to frequency judgment. That is, probability estimations can show greater average discrepancies from "reality" than do frequency estimates without necessarily producing a less appropriate pattern of estimates (i.e. lower correlations): such was the case for probe category #4. When probability estimation is inferior on both discrepancy and pattern measures, as it was for probe category #5, it is under conditions that produce relatively poor performance <u>generally</u>. A specification of these conditions is, however, beyond the scope of this study.

In summary, the following conclusions seem justified with respect to estimation performance: (a) Anticipation of a frequency or probability judgment task results in the use of similar strategies for encoding event occurrences and, therefore, attempts to induce a response set specific to either task is ineffective; (b) qualitatively different responses are produced as a function of the type of judgment elicited--frequency estimates are consistently superior to probability estimates; and, (c) the veridicality of both frequency and probability judgments is affected by the way in which events are defined (and the consequent range of absolute frequencies represented).

Predictive Choice. The accuracy of predictive choices was based on 80 choice trials administered during the fourth session. It will be recalled (Table 2) that the comparison was among five groups: the four experimental groups, which had previously made frequency or probability judgments congruent or incongruent with the response set, and a control group, which made neither type of estimation (and hence received neither estimation set). The 80 problems consisted of 60 binary choice pairs and 20 three-alternative choices, with both sets derived from an identical event domain. Thus it was possible to compare the performance of all five groups under these two choice conditions.

The accuracy of choices with a correction for guessing is shown along with the overall estimation performance in Table 8.

Table 8 about here

This adjustment was necessary for a valid comparison since probability of guessing correctly was different under the two choice conditions.

An analysis of variance revealed a significant group effect, $\underline{F}(4,40) = 4.91$, $p < .002$, choice condition effect, $\underline{F}(1,40) = 11.38$, $p < .001$, and group X choice condition interaction, $\underline{F}(4,40) = 3.10$, $p < .030$. The estimation groups were consistently better than the control group (see Table 8) thereby substantiating our earlier conclusion that estimations improve subsequent choice performance (Howell and Kerkar, 1981). Since the main thrust of the present study was to elucidate how this influence operates, however, our primary interest was in the comparison of the four experimental conditions.

The two most obvious ways in which prior estimation could affect subsequent choice would be through _general cuing_ (i.e. alerting the subject to the importance of probability/frequency information), and _summary encoding_ (i.e. providing a specific "memorial record" for use in the choice task). Were general cuing the principal mechanism, choice performance should be comparable for all four estimation groups; were summary encoding the principal mechanism, choice performance should reflect observed differences in quality of estimation (i.e. FF and PF should be superior to PP and FP).

Looking at Table 8, it is apparent that neither of these simple explanations is sufficient to account for this data. Estimation conditions <u>do</u> make a difference, ruling out the general cuing hypothesis as a complete explanation, and frequency estimates (FF and PF) <u>do not</u> yield better choices than probability estimates (PP and FP), ruling out the summary encoding hypothesis. What does seem to matter is the <u>congruity</u> of prior set and estimation requirement--a variable that had no appreciable influence on the quality of estimation performance. This totally unexpected result requires a considerably more involved (and admittedly post-hoc) account of the estimation-choice relationship. It would appear that the controlling factor may be something akin to task clarity and the resulting level of confidence attached to information used in the decision process. To illustrate, in the case of <u>congruent</u> estimation conditions, the subject benefits from general cuing, but he also may approach the choice task with confidence that the cued information (probability or frequency) is indeed relevant. By contrast, the <u>incongruent</u> estimation requirement may lead him to doubt the appropriateness of the information that he has been accumulating: even though he faces the choice situation with similar stored information, he may be less confident in it and therefore less inclined to follow its dictates.

This explanation underscores the importance of the subjects' task perception or their "task representation" in handling a decision problem (Einhorn and Hogarth, 1981). The idea that decision performance may be less than optimal due to a misperception of the task has been shown in other contexts (Lichtenstein and Feeney, 1968). The implication is that estimations lead to the development of an evaluative strategy for predictive choices. However, the incongruence produced in the PF and FP groups influences the task representation of the subjects and makes them less confident of using the same strategy consistently in attacking the newly encountered prediction problem. Moreover the lack of immediate feedback on the choice trials might deter the formulation and direct testing of rules. This is obviously not the case in the congruent condition--the FF and PP groups--and cuing leads to the use of an appropriate rule for choices.

To summarize, estimation serves a general cuing function in directing the subjects' mode of approaching the choice task. This accounts for the superiority of all the experimental groups over the control group. Furthermore, the development of a consistent strategy is subject to disrupting influences from changed perceptions of the task and this directly affects the accuracy of choices within the experimental conditions.

Turning to the other statistically reliable findings, significant differences appeared in the binary vs. triple choice conditions and the group X choice-condition interaction. Recall that choice pairs in the binary condition were derived from the other condition. Thus it was hypothesized that if the subjects develop a constant strategy, both choice conditions should produce equal levels of accuracy. Otherwise the three-alternative condition would compound uncertainty and the increased level of perceived complexity would be detrimental to performance. Although an interaction was expected, it was on the basis of the postulated effect of estimation task on choices. Thus the biasing effect of probability response was predicted to be augmented under the triple choice condition unlike the relatively veridical frequency response, which would produce comparable levels of performance for the two conditions. The control group was hypothesized to show an even greater decrease in accuracy compared to the experimental groups. However, the rather unexpected choice performance of the experimental groups and the explanation we proposed would lead to a different pattern of interaction.

The FF and PP groups should maintain the same level of performance on both choice conditions to the extent they adopt the same strategy consistently. Since the incongruence in the PF and FP groups disrupts their mode of dealing with the choice task, it should produce worse performance under the triple choice condition. The results provide some support for this possibility, especially the PF group whose accuracy drops considerably in the predicted direction. The other experimental groups maintain the same level of performance. Nevertheless, at this juncture great caution should be exercised in concluding too much from these results because of the variability in performance. One notable exception is the control group whose accuracy drops remarkably (42.5% and 26.7%). This means that when the subjects lack a proper strategy for dealing with uncertainty, increasing the number of alternatives compounds the problem further and leads to a deterioration of performance.

One interesting trend may be seen in the analysis of confidence judgments. The control group exhibited far more uncertainty in the accuracy of their choices than did any of the experimental groups. They tended to classify a larger percentage of their correct responses in categories denoting low confidence (randomly guessing or not very certain) than the other groups. This was statistically supported by a significant group X degree of confidence interaction obtained on an analysis of variance, $F(16,160) = 2.06$, $p < .02$. Whether this occurs as a result of the subjects' awareness in their own level of performance, where the lower accuracy of the control group was related to their low confidence or whether it results from confidence in the discovery of a rule (which the control group appeared to lack) is not clear.

CONCLUSION

The present findings substantiate our previous conclusion that predictive choice performance benefits from prior estimations. Since we attempted to clarify this by varying the quality of estimations and observing the effect on choice accuracy, any discussion of predictive choice behavior is contingent upon our conclusions about estimation performance.

In the case of estimations, set proved to be a nonsignificant factor, but the type of response elicited produced a reliable difference among groups. Thus anticipation of frequency or probability judgment did not lead to different kinds of encoding and performance was primarily a function of the nature of the elicited response. Frequency estimation produced a more veridical record of event occurrences than did probability estimation. This is consistent with our argument that the probability response entails an inherent bias and causes reliance on factors other than stored impressions of past event occurrencies. This bias, although pervasive, varied in magnitude as a function of the type of event category probed. It was also apparent that there were differences in the encoding of various dimensions and certain event categories produced better frequency and probability judgments than others. This aspect was, however, of only secondary interest and it was not investigated or discussed in depth.

Turning to predictive accuracy, the congruent set/estimation groups made the best choices followed by the incongruent groups and the control group performed the worst. This suggests that estimations provide the subjects with a strategy for dealing with the uncertainty entailed in choices. Furthermore, since the subjects' mode of attacking any decision problem essentially occurs in the context of how he perceives the task, an incongruity in this task representation deters the effective use of a single, consistent rule. However the subjects in the incongruent condition did exhibit an advantage over the control group and this argues strongly for the powerful cuing effect produced by the estimations.

Another interesting finding was the consistency of performance demonstrated by the experimental groups relative to the control under both choice conditions. The fact that increasing the number of alternatives might change the nature of the choice process has been suggested in other decision contexts (Payne and Braunstein, 1978). In the present study it served as a useful vehicle for compounding uncertainty in groups who lacked the "correct" approach to the choice task.

Our conclusions with regard to the determinants of choices are admittedly ad hoc and need further substantiation. The findings, however, have important practical and theoretical implications. They suggest that, contrary to normative predictions, choice behavior is dictated by other factors. Although judgment and choice are related, they are not synonymous (Einhorn & Hogarth, 1981; Einhorn, Kleinmuntz & Kleinmuntz, 1979). Future research should be directed at investigating the circumstances in which they can be identical, if at all. This might help us gain a better understanding of choice behavior and develop ways to promote better decision making.

Reference Notes

1. Hammond, K. R. *The integration of research in judgment and decision theory*. Center for Research on Judgment and Policy Report No. 226. Institute of Behavioral Science, University of Colorado, 1980.

2. Howell, W. C., & Kerkar, S. P. *Uncertainty measurement in a complex task as a function of response mode and event type characteristics*. Technical Report No. 81-1, Department of Psychology Report Series, Houston: Rice University, 1981.

3. Marques, T. E., & Howell, W. C. *Intuitive frequency judgments as a function of prior expectations, observed evidence, and individual processing strategies*. Technical Report No. 79-06, Department of Psychology Report Series, Houston: Rice University, 1979.

4. Marques, T. E. *Attention control in the frequentistic processing of multidimensional event streams*. Technical Report No. 80-13, Department of Psychology Report Series, Houston: Rice University, 1980.

5. Wallsten, T. *Using a conjoint measurement model to develop theory about probabilistic information processing*. Psychometric Laboratory Report No. 127. Chapel Hill, NC: University of North Carolina, 1975.

References

Asch, S. E. Forming impression of personality. Journal of Abnormal and Social Psychology, 1946, 41, 258-290.

Dawes, R. M. The mind, the model and the task. In H. L. Castellen and F. Restle (Eds.), Proceedings of the Seventh Annual Indiana Theoretical and Cognitive Psychology Conference, 1975.

Einhorn, H. J. The use of nonlinear, noncompensatory models in decision making. Psychological Bulletin, 1970, 73(3), 221-230.

Einhorn, H. J. 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. & Hogarth, R. M. Behavioral decision theory: Processes of judgment and choice. Annual Review of Psychology, 1981, 32, 53-88.

Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. Linear regression and process-tracing models of judgment. Psychological Review, 1979, 86(5), 465-485.

Hasher, L., & Chromiak, W. The processing of frequency information: An automatic mechanism? Journal of Verbal Learning and Verbal Behavior, 1977, 16, 173-184.

Hasher, L., & Zacks, R. T. Automatic and effortful processing in memory. Journal of Experimental Psychology: General, 1979, 108, 356-388.

Helson, H. Adaptation-level as a basis for a quantitative theory of frames of reference. _Psychological Review_, 1948, _55_, 297-313.

Helson, H. Current trends and issues in adaptation-level theory. _American Psychologist_, 1964, _19_, 26-38.

Hogarth, R. M. Cognitive processes and the assessment of subjective probability distributions. _Journal of American Statistical Association_, 1975, _70_(350), 271-294.

Howell, W. C. Storage of events and event frequencies: A comparison of two paradigms in memory. _Journal of Experimental Psychology_, 1973, _98_(2), 260-263.

Howell, W. C., & Burnett, S. A. Uncertainty measurement: A cognitive taxonomy. _Organizational Behavior and Human Performance_, 1978, _22_, 45-68.

Lichtenstein, S., & Feeney, S. J. The importance of data generating model in probability estimation. _Organizational Behavior and Human Performance_, 1968, _3_, 62-67.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. Calibration of probabilities: The state of the art. In H. Jungermann and G. de Zeeuw (Eds.), _Proceedings of the Fifth Research Conference on Subjective Probability, Utility and Decision Making_. Dordrecht, the Netherlands: D. Reidel, 1977.

Luchins, A. S. Mechanization in problem solving. _Psychological Monographs_, 1942, _54_(248).

Mertz, W. H., & Doherty, M. E. The influence of task characteristics on the strategies of cue combination. Organizational Behavior and Human Performance, 1974, 12, 196-216.

Payne, J. W., & Braunstein, M. L. Risky choice: An examination of information acquisition behavior. Memory and Cognition, 1978, 6(5), 554-561.

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

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

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.

Slovic, P., & MacPhillamy, D. Dimensional commensurability and cue utilization in comparative judgment. Organizational Behavior and Human Performance, 1974, 11, 172-194.

Tversky, B. Encoding processes in recognition and recall. Cognitive Psychology, 1973, 5, 275-287.

Tversky, B., & Teiffer, E. Development of strategies for recall and recognition. Developmental Psychology, 1976, 12(5), 406-410.

Wallsten, T. S. (Ed.) Cognitive processes in choice and decision behavior. Hillsdale, N. J.: Erlbaum, 1980.

TABLE 1

Distribution of emergency calls over 36 event categories classified by location, type of emergency and level of veracity.

	Type of Emergency			
	Police		Fire	
Location	AE	FA	AE	FA
1	4	4	1	1
2	2	0	1	0
3	0	0	0	2
4	2	2	10	0
5	8	2	1	0
6	0	1	8	0
7	4	4	0	0
8	8	2	0	0
9	0	4	2	2

Note: AE = actual emergency; FA = false alarm

TABLE 2

Description of the Conditions Constituting the Five Groups as a result of the manipulation of Response Set and Estimation Task.

Group	Response Set	Estimation Task	Designation
1	Frequency	Frequency	FF*
2	Frequency	Probability	FP**
3	Probability	Frequency	PF**
4	Probability	Probability	PP*
5	---	---	Control

Note: * Represents the congruent condition.

** Represents the incongruent condition.

TABLE 3

Illustration of the Temporal Sequence of Tasks.

	DAY 1		DAY 2		
Session #	1	2	3		4
Task	Dispatching (Acquisition)	Dispatching (Acquisition)	Dispatching (Acquisition)	Estimation for all groups except the "control" group	Predictive Choice

TABLE 4

Illustrations of Items in the Five Categories of Information Probed for the Two Estimation Tasks (Frequency and Probability).

Probe Category	Example Frequency Estimation (FF and PF groups)	 Probability Estimation (FP and PP groups)
1. Type of Event	How many total police calls did you receive?	If a call comes in, what are the chances (0-100%) that it will be a police call?
2. Type of Veracity	How many total false alarms did you receive?	If a call comes in, what are the chances that it will be a false alarm?
3. Event Type by Veracity	How many false alarms were police calls?	Suppose a call was a police call. What are the chances of its being a false alarm?
4. Event type by Location	(Map presented with instructions to estimate totals for events indicated), e.g., police calls, sector 1.	(Map presented with instructions to estimate the chances, 0-100%, for events indicated), e.g. police call, sector 1.
5. Event Type by Location by Veracity	(Map presented). Please fill in totals for false alarms only for events indicated, e.g. fire call, sector 2.	(Map presented). Suppose a call was a fire call. What are the chances it would be a false alarm in the indicated location, e.g. sector 2?

TABLE 5

Mean Error Scores for the Four Experimental Groups on the Deviation Measures across all Probe Categories.

Group	Unsigned Error	% Error
FF	4.28	32.77
FP	16.68	272.40
PF	4.18	30.93
PP	13.51	207.87

TABLE 6

Frequency and Probability Estimation Performance on the Five Probe Items with respect to the Deviation Measures.

Probe Category	Frequency* Estimation	Probability** Estimation	Frequency Estimation	Probability Estimation
	Unsigned Error		% Error	
1. Type of Event	6.28	6.39	13.36	10.20
2. Type of Veracity	7.06	9.61	29.40	30.04
3. Event Type by Veracity	4.61	11.76	24.27	46.42
4. Event Type by Location	2.12	14.12	60.99	288.57
5. Event Type by Location by Veracity	1.07	33.59	31.23	825.46

NOTE: * Frequency Estimation refers to the performance of FF and PF groups.

** Probability Estimation refers to the performance of FP and PP groups.

TABLE 7

Mean Correlations[1] for the Frequency and Probability Estimation Groups on the Two Spatial Probes (#4,5)

	Group	
Probe Category	Frequency* Estimation	Probability** Estimation
4. Event Type by Location	.68	.71
5. Event Type by Location by Veracity	.48	.25

[1] n = 18

*Frequency Estimation refers to the performance of the FF and PF groups.

**Probability Estimation refers to the performance of the FP and PP groups.

TABLE 8

Predictive Choice Accuracy with a Correction for Guessing and Overall Estimation Performance for the Five Groups.

	Percentage Correct Choices		Estimation		
Group	Binary Choice	Triple Choice	Unsigned Error	% Error	Correlation
FF	65.56	60.74	4.28	32.77	.61
FP	56.94	53.33	16.68	272.40	.47
PF	59.17	48.89	4.18	30.93	.55
PP	62.50	65.93	13.51	207.87	.49
C	42.50	26.67	---	---	---

END

DATE
FILMED

11-8

DTIC

AD-A106 143   RICE UNIV HOUSTON TX DEPT OF PSYCHOLOGY   F/G 5/10
CHOOSING AMONG ALTERNATIVES WITH UNCERTAIN OUTCOMES: EFFECT OF --ETC(U)
AUG 81 S P KERKAR, W C HOWELL   N00014-78-C-0555
UNCLASSIFIED TR-81-2   NL

2 OF 2
AD
A106143

SUPPLEMENTARY
INFORMATION

END
DATE
FILMED
1 82
DTIC

SUPPLEMENTARY

INFORMATION

AD-A106 143

Erratum: Choosing Among Alternatives With Uncertain Outcomes: Effect of Prior Cuing and Estimation Requirements (Technical Report #81-2)

Our computation of predictive choice accuracy (see Table 8, p. 42) was in error due to an inappropriate adjustment in the probabilities of guessing correctly for the binary and triple choice conditions. The corrected percentages of choices (refer to columns 2 and 3) should read as follows:

Table 8

Group	Percentage Correct Choices: Binary Choice	Triple Choice	Estimation: Unsigned Error	% Error	Correlation
FF	54.07	55.83	4.28	32.77	.61
FP	42.59	47.50	16.68	272.40	.47
PF	45.56	42.50	4.18	30.93	.55
PP	50.00	61.67	13.51	207.87	.49
C	23.33	17.56	---	---	---

In the light of these data, our earlier conclusions are modified somewhat. Since an ANOVA on the predictive choice accuracy revealed a significant effect of groups, $F(4,40) = 4.77$, $p < .003$, our conclusions regarding the cuing effect of estimations on choices remain unchanged. However, both the effect of choice condition and the group X choice condition interaction failed to reach significance, $F(1,40) < 1$, and, $F(4,40) = 1.98$, $p < .120$, respectively. Thus, our earlier interpretation that presentation of three alternatives (3-choice condition) serves to compound uncertainty relative to the binary choice condition remains only a theoretical possibility and received little empirical support.

We regret any inconvenience that this error may have caused to our readers.

Shanta P. Kerkar and William C. Howell