SOME FUNCTIONS OF MEMORY IN PROBABILITY LEARNING AND CHOICE BEHAVIOR

W. K. Estes

ROCKEFELLER UNIVERSITY, NEW YORK, NEW YORK

I. Introduction ........................................... 2
II. An Observation-Transfer Paradigm for the Study of Predictive Behavior .... 3
III. The Informational Basis of Predictive Behavior ........................................ 6
    A. Experiment 1. Probability Learning with Homogeneous Blocks of Observation Trials ........................................... 7
    B. Experiment 2. Randomized Training Trials and Delayed Tests .................. 10
    C. Experiment 3. Observation Trials with a Common Losing Alternative .......... 12
    D. Experiment 4. Joint Variation of Stimulus Frequency and Outcome Probability ........................................... 15
    E. Experiment 5. Joint Variation of Stimulus Frequency and Outcome Probability with Overlapping Observation Pairs .......... 17
IV. Theoretical Interpretation of Probability Learning Series .......................... 20
    A. The Associative Coding Model .................................. 21
    B. Parameter Estimates and Predictions of Choice Probability ..................... 24
    C. Scale Properties of the Model .................................... 29
V. Differential Reward Learning Relative to Values of Reward Sets ................... 30
    A. Experiment 6. Variation in Range of Reward Values and Mean Differences within Acquisition Pairs under a Full Information Condition .......... 31
    B. Experiment 7. Variation in Mean Reward Differences within Acquisition Pairs under a Partial Information Condition .......... 35
    C. Experiment 8. Mean Reward Differences within Disjoint Training Pairs ........ 37
VI. Theoretical Interpretation of the Differential Reward Series ........................ 38
VII. General Discussion ................................................................ 42
References ............................................................................ 44

1The research reported here was supported primarily by USPHS grant MH16100 from the National Institute of Mental Health; the theoretical analyses and the preparation of this chapter were supported also by grant MH23878. I wish to acknowledge the contribution of Edith Skaar to the data collection and processing.
I. Introduction

One of the main paths to generality and elegance in the development of physical theory has been associated with a search for measurable properties of objects and events with respect to which physical laws take on especially simple forms. With the advent of experimental psychology it seemed natural to try the same strategy. We need not pause to dwell on the accomplishments pursuant to this effort in the area of sensory processes where three-quarters of a century of effort has eventuated in a major body of quantitative theory concerning psychological measurement and scaling of sensory attributes. Extended beyond the sensory laboratory into the domain of human decision making and choice, the same methods have led to the concepts of subjective probability and utility. And in this area, the work of a number of theorists, perhaps most notably that of Thurstone (1927, 1959) and Luce (1959), has demonstrated that in many choice situations human beings do indeed behave as though they had recoded objective quantities into psychological magnitudes having simple scale properties.

In this latter line of theorizing, however, form has tended to outrun content. Consequently, we find ourselves in possession of a body of mathematical theory impressive with respect to both elegance and depth (for example, Krantz, Luce, Suppes, & Tversky, 1971) but with almost no demonstrable relevance to research going on in the various areas in which psychologists study decision and choice.

A principal reason for this annoying gap is to be found, I would suggest, in the fact that investigators interested in demonstrating scale properties of human choice behavior have been little concerned with the psychological processes that give rise to the scale properties. We can readily answer questions as to whether particular scales exhibit transitivity or satisfy interval or ratio properties, but if asked where scales come from, we fall silent. Surely scales involved in most choice and decision behavior of adult human beings must be generated by some as yet unidentified learning processes.

The studies to be reported in this chapter represent a first step toward illuminating the problem. I shall focus attention on two types of data, one having to do with subjective probability of events, and the other with utility or reward value. In each case the first step will be to demonstrate that orderly patterns of behavior reflecting scale properties can be generated in controlled learning situations. The next step will be to analyze each situation from the standpoint of
current theories of memory and to attempt to bring out experimentally some specific aspects of storage and retrieval that are responsible for the development and maintenance of the orderly patterns of choice behavior associated with psychological scales.

The first group of studies to be reported arose from an interest in problems of probability learning, probability estimates, and subjective probability. In the psychological literature, the term subjective probability has arisen primarily in connection with research on bets, gambles, and risk, and represents an inference from observed choice behavior with little attention to the type of memory structure that might underlie it (Cohen, 1964; Luce & Suppes, 1965). In probability learning experiments, on the other hand, it has been possible to show that orderly predictive behavior, and in some cases orderly probability estimates, arise as a function of an individual's experience with a set of alternative events over a series of trials in which the event probabilities remain constant.

In this latter line of research, we know quite a bit about the course of learning but almost nothing about the scale properties of the resulting choice behavior. If, following experience over a series of trials with two events whose probabilities are .75 and .25, an individual comes to predict these events with relative frequencies that match the true probabilities, we can conclude that he has learned something about the probabilities of the events. But we have no basis for concluding that he has developed a memory structure having any properties of a scale of probability. To take the latter step, we should need at least to demonstrate that, following experience with a few pairs of events drawn from a larger set, the individual is able in effect to place the various events on a psychological scale in such a way that he can then exhibit successful predictions or can generate accurate probability estimates concerning any new pairs of events from the set that might confront him on test trials.

II. An Observation-Transfer Paradigm for the Study of Predictive Behavior

In order to examine the acquisition of information regarding event probabilities, we need to simplify the subject's decision problem to the point that his choices can be taken as direct indicators of his state of information. Then at a later stage we may be able to produce specifiable states of information and study the development of decision strategies on a known baseline. Further, it seems essential to
restructure the traditional probability learning situation so that subjects can be expected to understand that their task is to learn about probabilities. In the past, the standard practice has been either to give subjects no meaningful orientation or to mislead them concerning the nature of the task (Estes, 1964; Myers, 1976). The reason in part is that orderly acquisition data and asymptotic probability matching seemed to be best attainable when subjects were led to believe that their task was to solve a problem or to make psychological judgments which had nothing to do with probabilities. Thus, in most studies of human probability learning, we were in effect dealing with incidental learning. On the few occasions when subjects were instructed as to the probabilistic nature of the task, learning proved more rapid (Peterson & Ulehla, 1965; Rubinstein, 1959). But then a new problem intruded. Learning curves for a substantial proportion of subjects rose quickly to an asymptote of 100% prediction of the more frequently occurring trial outcome, an inconvenient result if one's purpose is to determine how limits of learning are related to differences in outcome probability.

In the present research we have attempted to circumvent these problems by combining a task situation in which subjects clearly understand that they are dealing with probabilities and an experimental design that permits us to monitor the subject's state of information regarding probabilities without disturbing the learning process.

One widely familiar and well-publicized activity with regard to which nearly everyone in our culture understands the role of probability is the public opinion poll. It seemed that we could capitalize on this familiarity by setting up an experimental situation which would simulate the operation of an opinion poll. Our subjects, who would take on the role of observers of the poll results, could be expected to come to us well instructed by extra-laboratory experience prior to the experiment and to bring to the task whatever habits and skills they might have developed outside of the laboratory for dealing with uncertain events.

In the studies to be reported, our subjects were told that they were to imagine that the computer which operated the experimental apparatus had been programmed to conduct an imaginary opinion poll according to exactly the same principles as those governing the Gallup poll and others with which they were familiar. Prior to each trial of the experiment, a hypothetical individual in the hypothetical population being sampled would be interrogated by the computer as to his preference with regard to a pair of alternatives (these might be
candidates for office or treatments, such as headache remedies—for brevity we shall refer to them simply as stimuli) and the results of the inquiry would be transmitted by way of a simulated television screen to the subject who acted as an observer. The subject understood that his task was solely to observe the results of a series of trials, attempting to form a mental impression of the relative likelihoods that different stimuli would be preferred by the individuals being sampled, and that he then would be tested on his ability to predict the results of further polls in which the stimuli were tested in new combinations.

The point of departure for our new experimental paradigm was the observation of Reber and Millward (1968) that they could speed up probability learning considerably by starting their subjects off with a block of trials in which they simply observed the occurrences of the events which were later to be predicted but without making any responses on these trials. This result might well have been anticipated on the basis of analyses of Thorndikian learning situations which demonstrated the excessive information processing load entailed by the usual procedure of requiring a subject both to make responses and to observe outcomes on each learning trial (Buchwald, 1969; Estes, 1969).

Proceeding from these observations, it seemed that we might hope to obtain a more sensitive index of the learning which goes on in probability learning experiments by means of a transfer design which separates the occasions on which the subjects obtain information by observation from the occasions on which they are tested. The design is illustrated in Table I for an experiment in which three pairs of stimulus alternatives, listed at the top, appear on observation trials and then enter into transfer tests in all possible pairwise combinations. The probability that stimulus $A_i$ is the winner (i.e., is reported to have been preferred over the other member of the pair) on any observation trial is customarily denoted by $\pi_i$. The $\pi$ values shown are those used in the first two experiments to be reported.

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESIGN OF EXPERIMENTS 1 AND 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observation pairs</th>
<th>$A_1$ vs. $A_2$</th>
<th>$A_3$ vs. $A_4$</th>
<th>$A_5$ vs. $A_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of win ($\pi$ value)</td>
<td>.62</td>
<td>.38</td>
<td>.58</td>
</tr>
</tbody>
</table>
During the training phase of the experiment, the subject has opportunity to observe the results of a number of simulated preference tests on $A_1$ versus $A_2$, a number on $A_3$ versus $A_4$, and a number on $A_5$ versus $A_6$. Now, in the case shown in Table I, $\pi_1$ would be .62 and $\pi_3$ would be .58. Over a series of trials the subject would have an opportunity to learn that $A_1$ was preferred over $A_2$ 62% of the time, and $A_3$ over $A_4$ 58% of the time. With enough experience he might well come to encode stimuli $A_1$ and $A_3$ as "winners" and if asked to predict the results of future polls, always predict that $A_1$ would be chosen over $A_2$ and always to predict that $A_3$ would be chosen over $A_4$. How then could we determine whether the subject had acquired any information with regard to the differential win probabilities of $A_1$ versus $A_3$? Our procedure is to ask the subject to predict the outcome of a new test in which $A_1$ and $A_3$ are pitted against each other. He can do so with greater than chance success only if he has stored in memory information about the corresponding $\pi$ values beyond simply the fact that both $A_1$ and $A_3$ have been winners on the average.

This tactic is employed systematically in all of the experiments to be reported. Each experiment is divided into training and test phases. During the training phase the subject is given the results of repeated simulated preference tests on the observation pairs, then during the block of test trials he is asked to predict the results of surveys in which these stimuli are paired in all possible ways.

III. The Informational Basis of Predictive Behavior

In the initial exploratory studies conducted with the observation-test design I wished, first, to appraise the speed and precision of probability learning under these conditions and, second, to narrow down the range of tenable hypotheses concerning the basis of predictive behavior in memory. With respect to the latter objective, I began with the assumption that an individual's predictive behavior is based on his state of information regarding event probabilities. The information need not be numerical, nor even verbal, in character. Rather, the representation in memory of an event probability may be conceived as a cumulation of the residual effects ("memory traces") of previous trial outcomes. Further, the development of such a representation in memory need not depend on counting or other heuristics, but may be the result of a learning process which proceeds automatically as a
function of observation trials and is reflected directly in predictive behavior unless masked by hypothesis-testing activities on the part of the learner.

A. EXPERIMENT 1. PROBABILITY LEARNING WITH HOMOGENEOUS BLOCKS OF OBSERVATION TRIALS

1. Method

The design of this experiment required the subject to acquire information concerning the preferences that a hypothetical population of voters would exhibit concerning three pairs of candidates and to make predictions concerning results of elections.

The subjects were given full information concerning the probabilistic nature of the situation and the fact that their sole task was to gain sufficient information concerning probabilities of winning and losing on the part of various candidates during the observation blocks to improve their predictions for the various possible pairs on test blocks.

The candidates were represented by initials—single letters displayed on an oscilloscope screen interfaced to a PDP-8/I computer. In this experiment the basic plan for each session involved three pairs of letters; for each pair there were preassigned probabilities (π values) that each of the two members would be the winner or loser on each observation trial. The probability combinations were .62–.38, .58–.42, and .54–.46.

In Part 1 of the experiment, nine subjects were each studied for a single session in which they received eight replications of a basic cycle, each cycle comprising 72 observation trials, 24 on each of the three pairs of alternatives, followed by 30 test trials. An observation trial began with the display on the oscilloscope screen of the two letters corresponding to one of the pairs of hypothetical candidates for 500 msec, for example, A B,

after which a row of tallies appeared following each of the two letters, for example, A III B I,

and remained in view for 750 msec; then the screen went blank for a 1-sec intertrial interval. The numbers of tallies that might appear
ranged from 1 through 8 and the subject understood that on each trial the candidate who received the larger number of tallies was to be regarded as the majority choice, that is the "winner" of an opinion poll conducted with the sample of potential voters. When the tallies appeared on each observation trial, the subject pronounced the initial of the winner.

The computer program governing the experimental routine prescribed in advance which of the two alternatives would be the winner on each observation trial; the winner always received 5, 6, 7, or 8 tallies and the loser 1, 2, 3, or 4 tallies, the only constraint being that each particular value appear equally often within the winner and loser categories.

Within any observation series, each of the three pairs of alternatives was assigned to a block of 24 consecutive trials, the blocks occurring in random orders from one cycle to the next. Within a block, the same pair of alternatives appeared on each trial, the winner of each trial being determined by a random number generator with the constraint that in a .62-.38 block $A_1$ and $A_2$ were designated as winners exactly 15 and 9 times, respectively; in a .58-.42 block, $A_3$ and $A_4$ were winners 14 and 10 times; and in a .54-.46 block, $A_5$ and $A_6$ were winners 13 and 11 times, respectively.

During the block of test trials which followed each 72-trial observation series, all possible pairs of the six candidates were tested, each pair appearing once in each of the two possible left–right orders on the screen, resulting in a block of 30 test trials. The test trials were subject paced. At the beginning of the trial, two alternatives were presented (for example, $A_1$ versus $A_4$) and remained on the screen until the subject indicated his choice by pressing the response key under one of the two stimuli. The choice and the response time were recorded by the computer but no immediate feedback was given the subject regarding the actual winner of the hypothetical election in which the two candidates were paired. However, in order to maintain motivation, at the end of each 30-trial test block, the subject was shown the number of times on which his predictions of winners agreed with those generated by the computer, which had a full knowledge of the true probabilities of winning and losing for the population being sampled in the simulated opinion polls and elections.

Subjects in all of the experiments to be reported were young adults, not necessarily students, who in most cases were recruited via newspaper advertisements and who were paid for their services.
2. Results

For a group of nine subjects run under the procedure just described, the first overall result was a surprisingly rapid rate of learning, especially considering that the differences in $\pi$ value between the members of observation pairs are rather small compared to those involved in most classical probability learning experiments. For convenience in assessing learning, we define as a correct response on any test trial a choice of the alternative with the higher $\pi$ value. In these terms, the learning curves in terms of proportion of correct responses per test block rise from the initial value near .5 to a level slightly above .7 by the beginning of the fifth test block, then remain constant over the remainder of the session, never rising above .72. This terminal level is far above probability matching, which would be .56 over all test combinations of $\pi$ values.

The test results on particular pairs of alternatives are summarized in Table II in the form of a paired-comparison table. The rows from top to bottom and columns from left to right correspond to the stimulus alternatives (candidates) in order of $\pi$ values. The value in each cell represents the percentage of test trials on which the row alternative was chosen over the column alternative and the marginal values in the right-hand column represent the average proportions of cases in which each alternative was chosen over all others with which it was paired on test trials. These values represent data pooled over the entire session. The values in the upper right quadrant of the table, representing proportions of choices of the three highest valued stimuli (the winners on the observation trials) over the three lowest

<table>
<thead>
<tr>
<th>$\pi$ Value</th>
<th>.62</th>
<th>.58</th>
<th>.54</th>
<th>.46</th>
<th>.42</th>
<th>.38</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>.62</td>
<td>-</td>
<td>59</td>
<td>72</td>
<td>79</td>
<td>83</td>
<td>89</td>
<td>76</td>
</tr>
<tr>
<td>.58</td>
<td>41</td>
<td>-</td>
<td>55</td>
<td>77</td>
<td>83</td>
<td>85</td>
<td>68</td>
</tr>
<tr>
<td>.54</td>
<td>28</td>
<td>45</td>
<td>-</td>
<td>80</td>
<td>80</td>
<td>84</td>
<td>64</td>
</tr>
<tr>
<td>.46</td>
<td>21</td>
<td>23</td>
<td>20</td>
<td>-</td>
<td>60</td>
<td>63</td>
<td>37</td>
</tr>
<tr>
<td>.42</td>
<td>17</td>
<td>17</td>
<td>20</td>
<td>40</td>
<td>-</td>
<td>51</td>
<td>29</td>
</tr>
<tr>
<td>.38</td>
<td>11</td>
<td>15</td>
<td>16</td>
<td>37</td>
<td>49</td>
<td>-</td>
<td>26</td>
</tr>
</tbody>
</table>

$^a$Percentage choice of row over column alternative.
valued (the losers) are uniformly high. Further, the marginal values in the right-hand column of the table line up nicely in the order of the $\pi$ values of the alternatives even though $\pi$ values of adjacent stimuli within the winner and loser sets differ from each other by only .04.

Clearly the subjects could not have achieved these results entirely by encoding the stimuli as "winners" and "losers," for their percentages of correct predictions were somewhat above chance even when winners were paired with winners and losers with losers on transfer tests.

B. EXPERIMENT 2. RANDOMIZED TRAINING TRIALS AND DELAYED TESTS

It seemed possible that the exceedingly efficient learning in Experiment 1 might have been attributable, in part, to the procedure of utilizing homogeneous blocks of observation trials on specific pairs of alternatives. This procedure might have enabled the subjects to do a certain amount of counting or encoding of event runs as units in order to facilitate their estimates of probabilities. Also, although the subjects' principal source of information concerning event probabilities was the observation trials, we had no way to determine whether the feedback given at the end of each test block might have had any influence on performance. The present experiment was designed to obtain evidence on both of these questions.

1. Method

The general procedures, the stimulus sets, and the $\pi$ values were all the same as in the preceding experiment. There were just two changes in design. First, on the observation trials the three pairs of alternatives were presented in a completely random order rather than being segregated into homogeneous blocks. With this change in procedure it appeared quite impossible for the subjects to utilize counting or to keep track of properties of the preceding sequences on the various individual pairs during a 72-trial observation block. Second, for half of the subjects in this experiment, no test trials were given until after the fourth 72-trial observation block. There were seven subjects in each group.

2. Results

The terminal overall level of correct responding on test trials was approximately .67 as compared to approximately .71 in the first
experiment. The rate of approach to this terminal level also was similar, with little increase in correct response level over the last half of the session. For subjects receiving test blocks after every 72-trial block the proportion of correct responses after the fourth observation cycle was .66, whereas, for the group that did not receive tests following the first three observation blocks, the corresponding value was .65. Thus it appears that the information given at the end of the test blocks had little effect on learning.

Three additional groups of 10 subjects were run under the same procedures, but with a different set of letters for each group, and with test trials after each observation block. In the pooled data for all 37 subjects run under this condition, the overall percentage of correct responses increased from an average of 62 over test blocks 1 and 2 to 64 over blocks 3 and 4 and then 66 over blocks 5 to 8, where again 56% would constitute probability matching.

But principal interest attaches to the paired-comparison matrix for the transfer tests, presented in Table III for the data pooled over all test blocks. These choice percentages increase quite uniformly across the rows and up the columns of the table in just the manner one would expect if the choices were based on a representation in memory of the positions of the stimuli on a scale of subjective probability. And, considering the considerable increase in difficulty of the task over that of Experiment 1, the subjects exhibit rather striking proficiency at acquiring information concerning event probabilities, even under circumstances contrived to rule out the use of the heuristic devices or strategies that may supplement sheer learning by observation under less strictly constrained experimental conditions.

The results of this first series of studies offer some support for our

<table>
<thead>
<tr>
<th>π Value</th>
<th>.62</th>
<th>.58</th>
<th>.54</th>
<th>.46</th>
<th>.42</th>
<th>.38</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>.62</td>
<td>57</td>
<td>64</td>
<td>69</td>
<td>78</td>
<td>75</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>.58</td>
<td>43</td>
<td>59</td>
<td>69</td>
<td>71</td>
<td>76</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>.54</td>
<td>36</td>
<td>41</td>
<td>58</td>
<td>65</td>
<td>66</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>.46</td>
<td>31</td>
<td>31</td>
<td>42</td>
<td>59</td>
<td>56</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>.42</td>
<td>22</td>
<td>29</td>
<td>35</td>
<td>41</td>
<td>48</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>.38</td>
<td>25</td>
<td>24</td>
<td>34</td>
<td>44</td>
<td>52</td>
<td>36</td>
<td></td>
</tr>
</tbody>
</table>
preliminary analysis. The slow and variable course of probability learning characteristic of studies using traditional procedures is evidently attributable to extraneous activities on the part of the subjects that mask the course of acquisition. When normal adults understand fully the probabilistic nature of the task and are motivated to base their choices directly on their states of information, we obtain a picture of rapid and precise probability learning, with only tens rather than hundreds of observation trials being required to establish significant discriminations of very small differences in probabilities of events.

Further, the data clearly exhibit properties signifying the development of a memory structure that functions as a scale of a psychological magnitude. Following observational experience with only three of the pairs that can be formed from a set of six alternatives, our subjects generate orderly predictive responses to all possible test pairs, and, when training has been given with randomized observation trials, predict as well on new test pairs as on those included in the observation series. This transfer pattern is just what one would expect if the subjects had been given information regarding the relative placement of the candidates on a probability scale.

C. EXPERIMENT 3. OBSERVATION TRIALS WITH A COMMON LOSING ALTERNATIVE

The results of the first two experiments, and especially the second, appear quite compatible with the idea that the learner is rather directly translating objective into subjective probabilities and performing transfer tests on the basis of the placement of the various alternatives on the subjective probability scale. Now we wish to begin to look at experimental manipulations that might be more sharply diagnostic with respect to alternative interpretations. The main alternative to be considered, on the basis of theoretical considerations of memory, is that the individual is actually storing frequency counts in memory and then converting these to probability estimates at the time of transfer tests. In this experiment we examined a situation that should be the equivalent of the one studied in Experiment 2 from the standpoint of a subjective probability model but that might raise some new problems for an information processing system based on the acquisition of frequency information.

For this purpose we introduced the design exhibited in Table IV. The general procedures are the same as those of Experiment 2 except that on observation trials each of four candidates $A_i$ ($i=1-4$) is paired
TABLE IV
DESIGN OF EXPERIMENT 3 (COMMON LOSER)

<table>
<thead>
<tr>
<th>Observation pairs</th>
<th>Observation pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$ vs. $CL$</td>
<td>$A_2$ vs. $CL$</td>
</tr>
<tr>
<td>Probability of win ($\pi$ value)</td>
<td>.75</td>
</tr>
</tbody>
</table>

with a common alternative candidate who, on the average, is a loser with respect to all of the others.

All subjects were given six cycles each comprising 96 observation trials followed by 30 test trials. The tests included pairings of each of the five candidates that had appeared on observation trials with each other and also pairings of these with a novel alternative (N), introduced to the subject as a new candidate about whom they had to make predictions even though he had not participated in the simulated preelection opinion poll.

On observation trials the two candidates for whom simulated opinion poll data were to be presented were displayed in a vertical arrangement on the CRT screen with the common loser (CL) always in the lower position, whereas on test trials, the members of the various pairs appeared in a horizontal arrangement with left–right orders counterbalanced. The detailed procedure on observation trials differed for two experimental conditions. Sixteen subjects were assigned to Condition W, in which they were required on each observation trial to pronounce the initial of the candidate who received more votes on the trial; 32 subjects were assigned to Condition A, in which they were required on each observation trial to pronounce the name of the candidate appearing in the upper position and then to state whether he won or lost on the given trial. These manipulations were introduced in an attempt to influence the combinations of events to which the subjects would attend and therefore those which, on a frequency hypothesis, they might be expected to encode and store in memory.

The data are assembled in Table V, again in the form of paired-comparison matrices, the upper and lower panels representing Groups W and A, respectively. The overall rate and course of learning were similar to those of Experiment 2 and, if we look only at the transfer tests involving various pairings of the form $A_iA_j$, the results also are similar. At first glance the differentiation among these candidates (shown in the upper left-hand portion of each of the
TABLE V
RESPONSES ON PAIRED-COMPARISON TESTS OF EXPERIMENT 3
(COMMON LOSER)

<table>
<thead>
<tr>
<th>Alternative</th>
<th>(A_1)</th>
<th>(A_2)</th>
<th>(A_3)</th>
<th>(A_4)</th>
<th>CL</th>
<th>(N)</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition W</td>
<td>(\pi) Value</td>
<td>.75</td>
<td>.67</td>
<td>.62</td>
<td>.58</td>
<td>.34</td>
<td>(-)</td>
</tr>
<tr>
<td>(A_1)</td>
<td></td>
<td>.75</td>
<td>55</td>
<td>61</td>
<td>61</td>
<td>88</td>
<td>65</td>
</tr>
<tr>
<td>(A_2)</td>
<td></td>
<td>.67</td>
<td>45</td>
<td>51</td>
<td>47</td>
<td>50</td>
<td>56</td>
</tr>
<tr>
<td>(A_3)</td>
<td></td>
<td>.62</td>
<td>40</td>
<td>49</td>
<td>47</td>
<td>54</td>
<td>55</td>
</tr>
<tr>
<td>(A_4)</td>
<td></td>
<td>.58</td>
<td>39</td>
<td>53</td>
<td>53</td>
<td>53</td>
<td>57</td>
</tr>
<tr>
<td>CL</td>
<td></td>
<td>.34</td>
<td>39</td>
<td>50</td>
<td>46</td>
<td>47</td>
<td>55</td>
</tr>
<tr>
<td>(N)</td>
<td></td>
<td>(-)</td>
<td>12</td>
<td>11</td>
<td>13</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

Condition A

<table>
<thead>
<tr>
<th>Alternative</th>
<th>(A_1)</th>
<th>(A_2)</th>
<th>(A_3)</th>
<th>(A_4)</th>
<th>CL</th>
<th>(N)</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1)</td>
<td></td>
<td>.75</td>
<td>52</td>
<td>54</td>
<td>58</td>
<td>73</td>
<td>87</td>
</tr>
<tr>
<td>(A_2)</td>
<td></td>
<td>.67</td>
<td>48</td>
<td>53</td>
<td>50</td>
<td>70</td>
<td>89</td>
</tr>
<tr>
<td>(A_3)</td>
<td></td>
<td>.62</td>
<td>46</td>
<td>47</td>
<td>52</td>
<td>68</td>
<td>90</td>
</tr>
<tr>
<td>(A_4)</td>
<td></td>
<td>.58</td>
<td>42</td>
<td>50</td>
<td>48</td>
<td>68</td>
<td>87</td>
</tr>
<tr>
<td>CL</td>
<td></td>
<td>.34</td>
<td>27</td>
<td>30</td>
<td>32</td>
<td>32</td>
<td>79</td>
</tr>
<tr>
<td>(N)</td>
<td></td>
<td>(-)</td>
<td>13</td>
<td>11</td>
<td>10</td>
<td>13</td>
<td>21</td>
</tr>
</tbody>
</table>

paired comparison matrices), seems rather small; however, in view of the \(\pi\) values used, larger differences could not be expected on the basis of such theories as the stimulus sampling model. One can readily confirm, for example, that none of these observed transfer proportions for any pair \(A_iA_j\) are far from the value \(\pi_i/(\pi_i + \pi_j)\), which would represent the analog of probability matching in a simple noncontingent situation.

The transfer results for the \(A_i\) alternatives pitted against the CL prove to depend strongly on the conditions of vocalization of observation trial outcomes, yielding choice proportions much below probability matching for Condition W (name the "winner" on the trial) but above, on the average, for Condition A (report whether the \(A\) alternative won or lost). The critical differentiating factor appears to be the extent to which the instructions led subjects to attend to, and therefore encode in memory, the trial outcomes for the CL during the observation series.

The results on tests pitting the various alternatives that had been presented on observation trials against the novel candidate also present some surprises. Since the subjects had no information concerning the novel candidate, one might have expected them to assign him an average position, that is one close to the adaptation level
value on a probability scale. The high proportion of choices of the \( A_i \) alternatives over the novel candidate are compatible with this idea, but the fact that the CL was chosen over the novel candidate with similarly high probabilities (even higher than for \( A_i \) candidates in Condition W and only slightly lower in Condition A) is hard to fit in with the idea that the subjects were performing in accord with a scale of subjective probabilities. Evidently the subjects were in a sense misled by the fact that candidate CL, though always a loser on the average with respect to each of the candidates paired with him on observation trials, had accumulated a larger total number of winning outcomes during the observation series than any one of the \( A_i \) candidates. Thus in these data, as well as those for \( A_i \) vs. CL, we have a hint that the subjects' test performance is based at least in part on frequency information as distinguished from a rationally constructed scale of subjective probabilities.

D. EXPERIMENT 4. JOINT VARIATION OF STIMULUS FREQUENCY AND OUTCOME PROBABILITY

The hint emerging from Experiment 3 that subjects' performance reflects memory for total frequencies of winning outcomes on given alternatives as well as estimates of outcome probabilities is not unprecedented. In a related study (Estes, 1976) even more direct evidence of a similar tendency appeared. That study involved an observation-transfer design similar to that of the experiments reported above but a slightly different task orientation. Further, there was a major departure from the design of Experiment 2 of the present investigation in that some of the observation pairs occurred more often than others so that, for example, a stimulus with a \( \pi \) value of .42 occurred twice as often during the observation series as a stimulus with a \( \pi \) value of .54; thus the former stimulus accrued a larger total number of winning outcomes even though it had the lower probability of winning on any observation trial. The data for a transfer test on this pair of alternatives yielded a choice probability of .64 on the part of the subjects for the stimulus that had occurred more often but with actually the lower \( \pi \) value. If the subjects' test performance was governed by a scale of subjective probabilities, then the scale was not related in a monotone fashion to objective probabilities but rather was grossly perturbed by the variations in stimulus frequency.

Owing to its central importance to the objectives of the present study, that experiment was replicated with the present task orienta-
tion and the design illustrated in Table VI. The rows of the table labeled Conditions 1 and 2 indicate the relative frequencies with which the observation pairs above occurred during the observation series; for example, in Condition 1 the pair $A_3$ versus $A_4$ occurred twice as often as the pair $A_5$ versus $A_6$. Sixteen subjects were assigned to each condition and all were given six cycles of 72 observations followed by 30 test trials under conditions otherwise identical to those of Experiment 2 of the present paper. The way in which stimulus frequency and outcome probability interact in this situation is brought out most strikingly in the comparison shown in Table VII. These data represent proportions of choices of alternatives $A_1$ and $A_2$ over each of the other four alternatives. In both the upper and lower parts of the table the columns represent the alternatives $A_3$–$A_6$ in decreasing order of π value. However, one can see at a glance that, quite contrary to the results of the first three experiments, the relative probabilities of winning outcomes associated with the alternatives are exceedingly poor predictors of subjects' choices. Rather, the pattern of test performance reflects closely to the assignment of the stimulus alternatives to the high (H) or low (L) frequency condition.

One can readily infer from this last result that if we were to summarize the full test data for this experiment in paired-comparison tables like those presented for Experiment 2, and with the rows and columns again ordered in terms of π values, we would not obtain a similarly orderly pattern of data in the table. However, the paired-comparison table does provide us some leverage on the problem of determining just how stimulus frequency and outcome probability interact. We need only seek a function of these two variables such that, when the alternatives are properly ordered in terms of the function, the orderly pattern is restored to the paired-comparison table.

**TABLE VI**

DESIGN OF EXPERIMENT 4: VARIATION IN STIMULUS FREQUENCIES AND OUTCOME PROBABILITIES

<table>
<thead>
<tr>
<th>Observation pairs</th>
<th>$A_1$ vs. $A_2$</th>
<th>$A_3$ vs. $A_4$</th>
<th>$A_3$ vs. $A_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of win</td>
<td>.62</td>
<td>.58</td>
<td>.42</td>
</tr>
<tr>
<td>Relative frequency of pair</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>
TABLE VII

INDEPENDENT EFFECTS OF HIGH (H) OR LOW (L) STIMULUS FREQUENCY AND WIN PROBABILITY ON TEST PERFORMANCE IN EXPERIMENT 4

<table>
<thead>
<tr>
<th>( \pi ) Value</th>
<th>Condition 1</th>
<th>Condition 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>( n )</td>
<td>0.58</td>
<td>0.54</td>
</tr>
<tr>
<td>Value</td>
<td>36</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>.62</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>58</td>
<td>31</td>
</tr>
</tbody>
</table>

A solution turns out to be engagingly simple. We need only multiply the \( \pi \) value for each stimulus times the total number of times that it appeared during the observation series, thus in effect computing the total number of winning outcomes observed for the given alternative. The paired-comparison tables for the two conditions, ordered in terms of this index, are presented in Table VIII; and one may see at once that we have indeed restored the orderly pattern, with choice percentages increasing across the rows and up the columns as uniformly as one could expect for data which have some inherent variability. Further, the trends are not only qualitatively similar but quantitatively in very close agreement with the data obtained from the corresponding experiment in the study utilizing simulated preference surveys (Estes, 1976). Thus it appears that this result is highly reliable and replicable across variations in task orientation. The combined results appear to support strongly the suggestion emerging from Experiment 3 to the effect that our subjects' predictive behavior is based primarily on memory for event frequencies rather than upon accurate estimates of outcome probabilities.

E. EXPERIMENT 5. JOINT VARIATION OF STIMULUS FREQUENCY AND OUTCOME PROBABILITY WITH OVERLAPPING OBSERVATION PAIRS

Those who are attached to the idea of man as a rational decision maker might raise a question about the previous experiment on the grounds that the experimental design may not provide a fully adequate opportunity for learners to gain the information they need to establish a subjective probability scale. The possible weak point is that none of the observation pairs overlapped, so that, for example,
TABLE VIII
RESPONSES ON PAIRED-COMPARISON TESTS OF EXPERIMENT 4a

<table>
<thead>
<tr>
<th>Condition 1</th>
<th>112</th>
<th>90</th>
<th>80</th>
<th>54</th>
<th>52</th>
<th>44</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>112</td>
<td></td>
<td>64</td>
<td>69</td>
<td>79</td>
<td>79</td>
<td>80</td>
<td>74</td>
</tr>
<tr>
<td>90</td>
<td>36</td>
<td></td>
<td>58</td>
<td>68</td>
<td>67</td>
<td>78</td>
<td>63</td>
</tr>
<tr>
<td>80</td>
<td>31</td>
<td>42</td>
<td></td>
<td>69</td>
<td>54</td>
<td>64</td>
<td>52</td>
</tr>
<tr>
<td>54</td>
<td>21</td>
<td>22</td>
<td>31</td>
<td></td>
<td>49</td>
<td>60</td>
<td>37</td>
</tr>
<tr>
<td>52</td>
<td>21</td>
<td>33</td>
<td>46</td>
<td>51</td>
<td></td>
<td>61</td>
<td>42</td>
</tr>
<tr>
<td>44</td>
<td>20</td>
<td>22</td>
<td>36</td>
<td>40</td>
<td>39</td>
<td></td>
<td>31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition 2</th>
<th>104</th>
<th>90</th>
<th>88</th>
<th>56</th>
<th>54</th>
<th>40</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>104</td>
<td></td>
<td>52</td>
<td>70</td>
<td>74</td>
<td>82</td>
<td>82</td>
<td>72</td>
</tr>
<tr>
<td>90</td>
<td>48</td>
<td></td>
<td>55</td>
<td>66</td>
<td>78</td>
<td>78</td>
<td>63</td>
</tr>
<tr>
<td>88</td>
<td>30</td>
<td>45</td>
<td></td>
<td>61</td>
<td>74</td>
<td>76</td>
<td>57</td>
</tr>
<tr>
<td>56</td>
<td>26</td>
<td>34</td>
<td>39</td>
<td></td>
<td>64</td>
<td>76</td>
<td>48</td>
</tr>
<tr>
<td>54</td>
<td>18</td>
<td>32</td>
<td>26</td>
<td>36</td>
<td></td>
<td>56</td>
<td>34</td>
</tr>
<tr>
<td>40</td>
<td>18</td>
<td>22</td>
<td>24</td>
<td>24</td>
<td>44</td>
<td></td>
<td>26</td>
</tr>
</tbody>
</table>

Entries ordered by total win frequency over observation trials.

The learner had an opportunity to observe the relative probabilities of winning and losing on the part of alternatives \( A_1 \) and \( A_2 \) relative to each other but had no observation trials in which either of these was paired with any other member of the set of alternatives. Perhaps the necessary condition for the establishment of a rational scale of subjective probabilities is that the learner's observations include pairings of each alternative with every other, either directly, or at least by way of intermediary links (for example, \( A_1 \) versus \( A_2 \) and \( A_2 \) versus \( A_3 \)). This possibility is investigated in the present experiment with procedures similar to those of Experiment 4 except for a change in design.

The present experiment included four stimulus alternatives and two \( \pi \) values, combined into observation pairs according to the following scheme:

\[
\begin{align*}
A_1 & \text{ vs. } A_2 & A_3 & \text{ vs. } A_4 & A_1 & \text{ vs. } A_4 \\
.58 & \quad .42 & .58 & \quad .42 & .58 & \quad .42
\end{align*}
\]

Now all of the alternatives are connected in the sense mentioned above. Considering \( A_2 \) versus \( A_3 \), for example, the observation series provides information on \( A_2 \) versus \( A_1 \), \( A_1 \) versus \( A_4 \), and \( A_4 \) versus \( A_3 \), thus providing the basis for placing all of the alternatives in their appropriate positions on a probability scale. Twenty-eight subjects
were all given the same procedure, which involved eight cycles of 72 observation trials on the three observation pairs (in random sequence) and 12 test trials, including all possible pairings of the four alternatives in each of the two left–right orders.

During the observation series, the three observation pairs occurred equally often, but owing to the overlap the individual alternatives did not occur with equal frequency. Within a 72-trial observation block, of the $\pi = .58$ alternatives, $A_1$ occurred together with a winning outcome 28 times and $A_3$ 14 times, whereas for the $\pi = .42$ alternatives, $A_4$ occurred 20 times and $A_2$ only 10. Thus we again will be in a position to determine any effects of the independent variation of stimulus frequency and outcome probability.

The full paired-comparison data pooled over all eight test blocks are presented in Table IX. A glance at the table reveals immediately that once again the subjects were strongly influenced by stimulus frequency independently of outcome probability. For example, on the test pair $A_1$ versus $A_3$, both having $\pi$ values of .58 but differing in frequency, we find 76% choices of $A_1$; and in the $A_2$ versus $A_4$ pair, both members having had $\pi$ values of .42 but differing in frequency, we observe 68% choice of $A_4$. But most spectacularly, when we consider the test pair $A_3$ versus $A_4$, in which $A_3$ had the higher $\pi$ value but the lower frequency, we observe only 41% choices of $A_3$. It may be remarked, further, that these results almost certainly do not reflect incompleteness of learning, for similar analysis performed on only the last two test blocks reveals the same pattern with, for example, the “aberrant” choice proportion for the critical $A_3$ versus $A_4$ pair being even further accentuated (38% choices of $A_3$). And finally, we observe that with the entries in Table IX

| TABLE IX
RESPONSES ON PAIRED-COMPARISON TESTS OF EXPERIMENT 5
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_1$</td>
<td>$A_4$</td>
<td>$A_3$</td>
<td>$A_2$</td>
<td>Average</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>$A_1$</td>
<td>28</td>
<td>20</td>
<td>14</td>
<td>10</td>
<td>72</td>
</tr>
<tr>
<td>$A_4$</td>
<td>20</td>
<td>37</td>
<td>59</td>
<td>68</td>
<td>55</td>
</tr>
<tr>
<td>$A_3$</td>
<td>14</td>
<td>24</td>
<td>41</td>
<td>–</td>
<td>69</td>
</tr>
<tr>
<td>$A_2$</td>
<td>10</td>
<td>22</td>
<td>32</td>
<td>31</td>
<td>28</td>
</tr>
</tbody>
</table>

*Varying stimulus frequency and outcome probability with overlapping observation pairs.*
ordered in terms of total win frequencies per 72-trial observation block, we obtain the same orderly pattern of values, increasing up the columns and across the rows, observed for the same analysis of Experiment 4, whereas the order would disappear if the alternatives were ordered only in terms of \( \pi \) values.

There seems no remaining doubt that the subjects do not perform like rational statisticians who construct subjective probability scales making the best use of available information, but rather perform on the basis of memory for frequencies of the events to which they have been attending during observation trials.

IV. Theoretical Interpretation of Probability Learning Series

The first-order results to be interpreted present a relatively clear picture. All of the conditions yield systematic choice behavior on transfer tests, and in all conditions when the observation pairs have been presented in random sequence (that is, all experiments except Experiment 1), choice behavior on transfer tests is fully as accurate for new pairs as for observation pairs.

Further, the obtained paired-comparison matrices take on a standard form in all cases, with the probability of a choice of any one stimulus over another generally increasing as a function of their separation with respect to the independent variable used to order the stimuli of the rows and columns of the matrix. However, we have found that the independent variable has to change with experimental conditions. The \( \pi \) value, that is the probability of a winning outcome on occurrences of an observation pair, suffices for Experiments 1, 2, and 3, but total frequency of winning outcomes over the observation series proves to be a more appropriate independent variable for Experiments 4 and 5.

With few exceptions differences in average paired-comparison values (the row means of Tables II, III, V, VIII, and IX) between different stimuli prove directly related to differences in \( \pi \) value or win frequency, as appropriate. One exception will be observed in the case of Experiment 1 (Table II) where a disproportionately large difference in average paired-comparison value appears between the stimuli with \( \pi \) values of .54 and .46. A second deviation occurs in Experiment 3 where the paired-comparison value for the CL is out of line with expectation on the basis either of \( \pi \) value or of overall win
frequency. A third discrepancy is the deviation from monotonicity in Experiment 4, Condition 1 (that is the inversion of the fourth and fifth row averages in the upper matrix of Table VIII).

The idea of a single subjective scale underlying predictive behavior in all of the conditions studied here can be maintained only if we assume that the discrepancies all represent noise in the data. That interpretation would be a possibility; but should we facilely accept the conclusion that deviations from a preconceived idea of regularity necessarily represent noise? I shall instead entertain the view that the discrepancies as well as the regularities are real. Proceeding on this assumption, I can find no single independent variable that straightens out all of the paired-comparison functions. However, ideas developed elsewhere (Estes, 1976) regarding the interpretation of probability learning in terms of concepts of memory suggest the possibility that the irregularities arise because the data represent, in effect, a mixture of choice behaviors based on two or more different psychological scales.

The reason for anticipating a mixture is that, depending on instructions, experimental context, and previous experience, different subjects may attend selectively to different events. In related studies (Estes, 1976) we have found that variations in task orientation led to different tendencies on the part of subjects to attend to and encode winning as compared to losing outcomes in situations like those studied in the present experiments. Further, faced with a similar problem in relation to subjects' ability to generate linear orderings on the basis of frequency information, Humphreys (1975) arrived at a hypothesis of a mixture of attentional tendencies toward correct and incorrect outcomes of verbal discrimination trials. Consequently, I propose now to seek an interpretation of all of the trends in the present data, both the regularities and the deviations already noted, in terms of a model predicated on the idea that probability judgments are based on memory for relative frequencies of the events to which subjects selectively attend during an observation series.

A. THE ASSOCIATIVE CODING MODEL

As applied to the present situation, we would assume on the basis of this model that test performance is mediated by categorical memory for relative frequencies of stimulus-outcome combinations. We conceive that the learner encodes the principal event categories to which he attends under a given condition, in this case stimuli to-
gether with winning or losing outcomes, and stores memory trace vectors that incorporate these encoded representations of event combinations together with contextual cues common to the observation and test situations. A trace vector may be denoted $T_{xAO}$, where $x$ refers to the context, $A$ the stimulus alternative and $O$ the outcome (W or L) of an observation trial. The contextual cues are assumed to vary in availability from trial to trial (as in standard stimulus sampling theory or encoding variability theory) so that, over a series of trials, the number of traces established that include particular event combinations will on the average be proportional to their relative frequencies of occurrence. Owing to different task orientations, subjects may vary from one experiment to another in their probabilities of attending to and encoding particular event categories, here winning or losing outcomes.

These ideas can be given a simple quantitative representation in the following way. When learning is complete, if the subject has been encoding only winning outcomes on observation trials, then the proportion of all memory trace vectors stored which are associated with winning outcomes for alternative $A_i$ can be represented by a quantity $\alpha_i$, where $\alpha_i$ equals $W_i/\Sigma W_k$. In this expression, $W_k$ denotes the actual frequency of wins over the observation series for alternative $A_k$ ($k = 1, 2, \ldots, N$, where $N$ is the number of alternatives). Similarly, if the subject were encoding only losing outcomes, then the proportion of trace vectors associated with losing outcomes for alternative $A_i$ would be $\beta_i = L_i/\Sigma L_k$, $L_k$ denoting the actual frequency of losses for $A_k$ over the observation series.

With regard to the way in which choices are generated on the basis of the information in memory, it is assumed that on any test trial, the subject scans the two test stimuli and attempts to recall their previously associated outcomes. On a test trial when, say, alternatives $A_i$ and $A_j$ are presented, we assume that the subject scans the alternatives, terminating with a choice when he finds a match between the pattern of a test stimulus plus available contextual cues and a memory trace stored on an observation trial when a particular type of outcome occurred. It can be shown that very simple expressions relate asymptotic choice probabilities to the quantities $\alpha_i$ and $\beta_i$ defined above, the form of dependence depending on the type of outcome encoded (for details, see Estes, 1976).

Letting $P_{ij}(i)$ denote the probability of choosing alternative $A_i$ when alternatives $A_i$ and $A_j$ are presented for choice, the predicted values of this choice probability are given by Eqs. (1), (2), and (3),
respectively, according as the subject has attended only to winning outcomes, only to losing outcomes, or to both, on observation trials:

\[ P_{ij}(i) = \frac{\alpha_i^2}{\alpha_i^2 + \alpha_j^2} \]  
(1)

\[ P_{ij}(i) = \frac{\beta_j^2}{\beta_i^2 + \beta_j^2} \]  
(2)

or

\[ P_{ij}(i) = \frac{\alpha_i \beta_j}{\alpha_i \beta_j + \alpha_j \beta_i} \]  
(3)

Since we do not in general know to what events subjects have been attending, and wish to allow for some mixture within a group of subjects, we introduce three parameters, \( w, x, \) and \( b \), to represent the proportions of instances in which Eqs. (1), (2), and (3), respectively, are applicable in a given experiment. Then by evaluating these parameters from the data, we can hope to determine what mixture of modes of selective attention is represented in the behavior of the subjects. Further, we must allow for the possibility that under particular experimental conditions learning may not be complete; to this end we introduce also a parameter \( \phi \) to represent the probability that learning has occurred in the case of any given subject and item (hence, with probability \( 1-\phi \) learning has not occurred and choices should be expected to conform to a chance probability of \( .5 \)).

Combining these various terms we arrive at the following equation:

\[ P_{ij}(i) = (1-\phi)(.5) + \phi \left[ w \frac{\alpha_i^2}{\alpha_i^2 + \alpha_j^2} + x \frac{\beta_j^2}{\beta_i^2 + \beta_j^2} + b \frac{\alpha_i \beta_j}{\alpha_i \beta_j + \alpha_j \beta_i} \right] \]

(4)

representing predicted choice probability in terms of a weighted combination of unlearned and learned states and, in the latter case, different proportions of instances in which winning outcomes, losing outcomes, or both have been attended to.

To evaluate the fruitfulness of this development in the present situation, I shall now proceed in two steps. The first is to determine
the best fit of Eq. (4) to the data of each of Experiments 1–5; the second is to consider the bearing of the results on the question of the nature of the psychological scales underlying the observed choice behavior.

B. PARAMETER ESTIMATES AND PREDICTIONS OF
CHOICE PROBABILITY

The first step was initiated by a computer search of possible combinations of parameter values to determine the best fit of Eq. (4) to the full paired-comparison matrices of Tables II, III, V, VIII, and IX according to the least-squares criterion. There are 15 independent observed proportions in the paired-comparison matrices for Experiments 1–4 and six in Experiment 5; hence, even though we have to estimate the values of four parameters, the number of degrees of freedom in the data substantially exceeds the number of parameters estimated. Further, since on a preliminary analysis the parameter $x$ turned out to be near zero in all cases, $x$ was set equal to zero, thus dropping the term $\beta_j^2/(\beta_i^2 + \beta_j^2)$ from Eq. (4) and reducing the number of free parameters to three. It appears that under the conditions studied here, there is very little likelihood that subjects will attend only to losing outcomes. That result was, however, obtained in an experiment reported elsewhere in which subjects were instructed to pronounce the names only of losing alternatives on observation trials (Estes, 1976). As may be seen from the parameter estimates summarized in Table X, it turns out that for all of the experiments in which the observation pairs were nonoverlapping, we are dealing with essentially asymptotic data ($\phi=1$).

Now I shall comment briefly on the problems encountered in treating the data of each of the experiments in turn in terms of the model. To provide comparability across experiments I have averaged the values associated with each stimulus (that is the values across the rows of the observed and predicted paired-comparison matrices) and plotted these as a function of the proportion of W's ($\alpha_i$ in terms of the model) in Figs. 1–4.

1. The first attempts to fit the data of Experiment 1 failed to yield a reasonable result, owing to the large gap between the mean paired-comparison values for the lowest three and highest three stimulus alternatives. Consideration of the special features of this experiment in terms of memory factors suggested that the key to the difficulty might lie in the fact that in this experiment, only, the observation
TABLE X
PARAMETER ESTIMATES FOR PROBABILITY LEARNING EXPERIMENTS

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Learning rate ($\phi$)</th>
<th>Weight of component reflecting attention to:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>W's ($w$)</td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0</td>
</tr>
<tr>
<td>3 Group W</td>
<td>.50</td>
<td>.37</td>
</tr>
<tr>
<td>3 Group A</td>
<td>.57</td>
<td>0</td>
</tr>
<tr>
<td>4 Cond. 1</td>
<td>1.00</td>
<td>.70</td>
</tr>
<tr>
<td>4 Cond. 2</td>
<td>1.00</td>
<td>.80</td>
</tr>
<tr>
<td>5</td>
<td>.95</td>
<td>.71</td>
</tr>
</tbody>
</table>

$^a$ For Experiments 1 and 2 the theoretical asymptotes for attention to W or W & L are almost indistinguishable, so the estimates of the weights are arbitrary.

trials were blocked, so that a particular pair of alternatives appeared on one block of 24 consecutive trials within each 72-trial cycle. Under this condition, it might have been especially easy for subjects to notice that one member of each observation pair was the winner on the average and thus to encode the two members of each pair as "winner" (W) and "loser" (L) or the equivalent. Then on all test trials in which a winner was paired with a loser the subject would be expected to choose the winner without having to consult his memory for relative frequencies of previous occurrences. On the supposition that this type of encoding occurred half of the time and that in the remaining instances choice behavior was governed by the relative frequency model with the parameter values carried over from Experiment 2 (discussed below), we arrived at the theoretical function shown in Panel 1 of Fig. 1. We appear to have quite a good description of the observed paired-comparison function with only one free parameter (the proportion of W-L encodings) estimated from these data. A similar conception of performance reflecting a mixture of memory for frequency and outcome encoding has been found useful by Medin (1974) in the interpretation of verbal discrimination data.

2. Since the proportion of correct responses (that is, choices of the member of a test pair which had actually the higher $\pi$ value) was
virtually constant over successive test blocks, it was clear that the data of this experiment represent essentially asymptotic performance, so the parameter $\phi$ in Eq. (4) was set equal to 1. In evaluating the weighting parameters, $w$, and $b$, a small problem arose. The first two experiments were the only ones in this series conducted before the ideas embodied in the relative frequency model began to play a role in planning the experimental designs, and as a consequence it happens that for Experiments 1 and 2 the predicted asymptotic choice probabilities for Eq. (1) and (3) are virtually identical. Although the estimate of $b = l$ was generated by the search procedure, a mixture of $w = .50$ and $b = .50$ yields an almost indistinguishable fit. The data for this experiment appear a bit noisy, owing no doubt to the very restricted range of $\pi$ values, but nonetheless the correspondence between theoretical and observed values is fairly satisfactory (Panel 2 of Fig. 1).

3. In Experiment 3, which involved the pairing of each of four different candidates with a "common loser" during the observation series, we obtain rather peculiar looking paired-comparison functions (Fig. 2) owing to the fact that the stimulus alternative which was on the average a loser in each of the observation pairs yields mean paired-comparison values distinctly lower than would be expected solely on the basis of the total proportion of winning outcomes accrued to this alternative. The procedure of requiring the subjects to pronounce the name of the alternative paired with CL and also to state whether the alternative won or lost on each observation trial (used with Group A) yields distinctly faster learning than the standard procedure of requiring subjects only to pronounce the name of the winner (used with Group W); the difference in learning rate is manifest in the higher estimated value of $\phi$ for the former group (Table X) and the fact that the paired-comparison curve is elevated considerably, although similar in form. The parameter estimates
Fig. 2. Mean paired-comparison scale values (connected closed circles and closed triangles) for Experiment 3. The subjects in Group W pronounced the name of the winner on each observation trial, whereas the subjects in Group A pronounced the name of the candidate paired with the common loser and stated whether he won or lost. The open circles for Group W and open triangles for Group A represent predicted values from the model.

Further reflect the difference in instructions in that the value of $b$ is appreciably larger for Group A than for Group W. With these parameter values, the model has no difficulty in accounting for the curiously shaped paired-comparison functions of both groups.

4. Again for Experiment 4, the data appear to represent asymptotic performance, with $\phi$ equal to 1 for both Conditions 1 and 2. The estimation procedure yields the parameter values shown in Table X for each condition. As expected, since the subjects were instructed to pronounce only the names of winning alternatives on observation trials in this experiment, the parameter values showed that the subjects attended only to W's in the majority of instances. The theoretical functions computed using these parameter values and plotted in Fig. 3 provided rather good accounts of observed functions which otherwise might have seemed rather aberrant in form. In particular, both the drop in paired-comparison value from the second to the third point in Condition 1 and the very abrupt increases from
the second to the third and fourth to the fifth points in Condition 2 are nicely handled by the model. The reason, in terms of the theory, for these deviations from smoothly increasing functions is that the independent variable in the figure is the proportion of winning outcomes associated with each stimulus alternative but the parameter estimates shown in Table X indicate that the subjects were on some occasions attending to both winning and losing outcomes. The third point for Condition 1 in Fig. 3 represents a stimulus alternative that had a higher proportion of winning outcomes than the one associated with the second point but also a disproportionately larger proportion of losing outcomes. Similarly, the sharp upturns in the curve for Condition 2 reflect points at which increases in proportion of winning outcomes were associated with substantial decreases in proportion of losing outcomes.

5. It will be recalled that in Experiment 5 the observation pairs were “connected” in the sense that for any two stimulus alternatives $A_i$ and $A_k$ which were not paired with each other during the observation series, each was paired with a common third alternative $A_j$. This design, together with instructions to pronounce only the names of winners on observation trials, yields an exceedingly orderly paired-comparison function and one closely described by the model (Fig. 4). As in the case of Experiment 4, the data appear to be asymptotic and the parameter estimates indicate that the subjects attended only to winning outcomes in a large majority of cases, but with a small admixture of instances in which they attended to both W’s and L’s.

Taking these results together, it appears that a model which assumes the basis of probability judgments to lie in memory for relative frequency information accounts both for the orderly scale-like choice performance that appears under optimal conditions and

![Fig. 4. Mean paired-comparison values for Experiment 5 (connected closed circles) together with predicted values (open circles).](image)
also for the perturbations, sometimes large, that appear under various unusual circumstances.

C. SCALE PROPERTIES OF THE MODEL

Now to take the next step in our program of analysis and consider more formally the relation of the results to the conception of a psychological scale underlying choice, we return to Eqs. (1)–(3), which represent the asymptotic choice probabilities predicted by the model for the several pure cases of selective attention. We can convert these expressions to the form shown in Eqs. (5)–(7) by dividing the numerator and denominator of the right side of each equation by the numerator, obtaining

\[ P_{ij}(i) = \frac{1}{1 + (\alpha_j^2 / \alpha_i^2)} \]  
(5)

\[ P_{ij}(i) = \frac{1}{1 + (\beta_i^2 / \beta_j^2)} \]  
(6)

and

\[ P_{ij}(i) = \frac{1}{1 + (\beta_i / \alpha_i)(\alpha_j / \beta_j)}. \]  
(7)

It will be noted that each of these turns out to be of the form

\[ P_{ij}(i) = \frac{1}{1 + (\nu_j / \nu_i)} \]  
(8)

where \( \nu_k = \alpha_k^2, 1/\beta_k^2 \), and \( \alpha_k / \beta_k \), for \( k = i, j \), in the case of Eqs. (5)–(7), respectively. This common form will be recognized to be precisely that used to predict paired-comparison choices on the basis of a model assuming that the values of the alternatives fall on a ratio scale of some psychological magnitude (Luce, 1959).

Thus our result indicates that we can regard our subjects as making choices, at the asymptote of learning, on the basis of the relative positions of the stimulus alternatives on psychological scales which have their bases in memory representations of relative frequency information. The reason we could not have made sense of our data if we had started from a scaling approach is that we turn out to have a mixture of cases in which subjects have been selectively attending only to winning outcomes and thus are operating on one scale and
cases in which they have been attending to both wins and losses and thus are operating on another.

Since a meticulous analysis of the situation in terms of concepts of memory has been necessary to bring order out of the data, one might ask what is added by introducing the notion of a psychological scale. The answer would seem to be that there is no immediate gain with regard to predictability of choice behavior in this situation. However, there may be some long-term benefits by way of bringing out communalities between choice behavior in learning situations and choice behavior as it has been dealt with traditionally in studies of preference in the scaling tradition. In particular, the concept of a scale may help to bring out communalities between probability learning and multiple choice learning, to which we turn in the next section.

V. Differential Reward Learning Relative to Values of Reward Sets

The preceding series of experiments shows clearly enough that in a probability learning situation subjects do develop a memory structure having properties of a psychological scale, though evidently the structure represents memory for relative frequencies of events rather than subjective probabilities as usually conceived. Now we wish to see whether a similar process operates in situations where the subject's problem is learning of relative reward magnitudes.

Much of our information concerning the functions of information and incentives in simple human learning comes from experiments that utilize variations on the standard multiple choice situation (see, for example, Estes, 1966; Keller, Cole, Burke, & Estes, 1965). Typically the task involves a list of items, each item comprising a pair of stimuli with associated reward values that are initially unknown to the subject. On a series of trials items occur in random order; as each item occurs, the two stimuli are presented, the subject makes a choice between them and then is given information either concerning the reward associated with each of the stimuli (full information condition) or only concerning the reward given for the alternative he chose (partial information condition).

The results of the studies just cited and others that I have reviewed elsewhere (Estes, 1969) lean heavily in favor of a cognitive interpretation of adult human learning in this type of situation. A simplified cognitive model would assume that the subject simply
learns in paired-associate fashion the relations between stimuli and reward values and then bases his decisions on this information. However, studies by Allen (1972) and Allen and Estes (1972) have indicated that the full picture may not be quite so simple. Evidence was obtained that subjects often solve problems in the sense of meeting a criterion of 100% correct choices of the higher-valued alternative without being able to recall the explicit values when tested by means of a memory probe. A possible interpretation suggested was that, rather than attempting to learn explicit stimulus-reward associations, subjects may attempt to rehearse and encode labels for stimuli which they observe to be followed by high reward values. This differential rehearsal would, then, provide a possible mechanism for generating representations of relative frequency information in memory comparable to those inferred from the data of the probability learning experiments discussed in the previous sections of this chapter.

A. EXPERIMENT 6. VARIATION IN RANGE OF REWARD VALUES AND MEAN DIFFERENCES WITHIN ACQUISITION PAIRS UNDER A FULL INFORMATION CONDITION

If the frequency coding hypothesis is sound, then it should be possible for subjects to learn multiple-choice problems even under conditions so arranged that the learning of explicit stimulus-reward associations is impractical. In the present experiment, and those following in this series, an attempt was made to produce the desired conditions by using a design and procedure generally similar to those of previous multiple-choice experiments except that, in making up the problems, each stimulus was associated with a set of reward values rather than with a single value, and a probability distribution was defined over the set of possible values.

For example, in Problem 1 of the present experiment, one stimulus was associated with the reward values 6, 8, and 9 and a second stimulus with the values 4, 6, and 9. In each case the values were equiprobable so that if these two stimuli were presented for choice the subject would, on the average, obtain a larger reward by choosing the first stimulus over the second.

Four problems were constructed, each consisting of a set of five stimulus alternatives. We shall denote these $A_i$, but as displayed to the subjects they were represented by random sets of five letters of the alphabet, each stimulus being associated with a set of three
reward values. All pairs that could be made up from the set of five stimuli were utilized in the course of a subject's experience on a given problem. The general idea was that a subset of four of these pairs would be presented during the acquisition phase; then the subject would be tested on all of the remaining pairs that could be made up from the given set of stimuli.

As illustrated in Table XI, Problems 1 and 2 were identical with respect to the average reward values of the five stimuli, as were Problems 3 and 4; and Problems 1 and 3 were identical with respect to the particular stimulus pairs presented on acquisition trials, as were Problems 2 and 4. As in the experiments on predictive behavior we wish to determine whether experience with a subset of the pairs that can be formed from a set of alternatives will permit the subject to generate a memory structure which can mediate correct performance on other test pairs in the manner that would be expected if his memory structure has the properties of a psychological scale. With the design illustrated in Table XI, we will be able to determine the way in which the acquisition of information reflects both the range of reward values and the mean reward differences within the stimulus pairs that are presented on the acquisition trials.

In the experimental situation the subject was seated in front of a teletypewriter which operated under the control of a PDP-8 I computer. Material could be typed out on the typewriter under the control of the computer, and when required control could be

### TABLE XI

**DESIGN OF EXPERIMENT 6**

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>Av. reward values</th>
<th>Acq. pairs&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Problems 1,2</td>
<td>Problems 3,4</td>
</tr>
<tr>
<td>A₁</td>
<td>7.7</td>
<td>6.3</td>
</tr>
<tr>
<td>A₂</td>
<td>6.3</td>
<td>5.7</td>
</tr>
<tr>
<td>A₃</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>A₄</td>
<td>3.7</td>
<td>4.3</td>
</tr>
<tr>
<td>Aₛ</td>
<td>2.3</td>
<td>3.7</td>
</tr>
</tbody>
</table>

<sup>a</sup>Varying mean reward differences with full information.

<sup>b</sup>‘x’s in each column denote the alternatives of training pairs, e.g., for Problems 1,3 A₁ vs. A₂, A₂ vs. A₃, etc.
switched to the subject, who could communicate his responses back to the computer by typing them on the keyboard. The subject was told that he was to play a game against the computer, which operated according to certain specific rules. The game would consist of a series of trials on which he would be presented with pairs of stimuli and would be rewarded for his choices. The rewards would take the form of point values which at the end of the session would be converted into monetary payoffs. The specific rule under which the computer operated varied from experiment to experiment.

Subjects were run on one problem at a time. On each problem the acquisition series consisted of 60 trials, a block of 10 trials on each pair followed by a block of five trials on each pair. The transfer series consisted of a block of five trials on each of the six remaining pairs, the reward assignments and procedure being unchanged from acquisition to transfer trials. Each trial of the experiment began with presentation of two stimuli, for example:

A or B?

Then when the subject typed in his choice, say B in the example, the reward sets associated with the two stimuli were typed out by the teletypewriter below them on the display, which might now look as follows:

A or B?
B 689 257

Finally, the reward which the subject received for his choice, 5 in the example and the cumulative total points received through the given trial were typed out.

Reward = 5 Total = 61

After a delay of 1–2 sec the stimuli for the next trial were presented and so on.

The subjects were sixteen young adults all of whom were run on all four problems except for a few instances of incomplete sessions owing to apparatus or scheduling problems.

The first conspicuous result with this procedure is that even though the acquisition conditions made it impractical for the subjects to learn associations between stimuli and specific reward values, they did acquire information on the basis of experience with a limited number of stimulus pairs that permitted them to choose with substantially better than chance success on new test pairs. If we look
only at the very first trial on which each test pair appeared, the average proportions of correct choices (that is, choices of the alternative with the higher average reward value) were .82, .66, .62, and .62 for Problems 1–4, respectively.

A better picture of the quantitative properties of test performance can be obtained by pooling data over the five transfer trials and assembling these choice percentages in paired-comparison tables for Problems 1 and 2 and for Problems 3 and 4 combined, as displayed in Table XII. It is apparent that, in spite of the smaller number of observations in the present experiment, the paired-comparison data yield the same orderly pattern that we saw repeatedly in the case of the experiments on predictive behavior. To obtain an adequate basis for comparing the effects of the variations in conditions that differentiated the four problems, I have combined the choice proportions for the terminal acquisition block on each problem and the full five test trials on each test pair into a single paired-comparison table and then determined a scale value for each stimulus by computing the proportion of times that it was chosen over all of the other stimuli of the set. This procedure yields values corresponding to the row averages in Table XII and other preceding paired-comparison tables [and essentially the equivalent of the scale values obtained by Thurstone's Case IV in classical paired comparison scaling (Thurstone, 1927)]. The values so obtained for the four problems are plotted in Fig. 5. The two striking features of this figure are the substantial difference between the first two and second two problems, reflecting the difference in range of reward values during acquisition, but the very close similarity between Problems 1 and 2 and between Problems 3 and 4, indicating that the memory structure developed by the subjects is quite independent of the particular subset of stimulus pairs with which they had experience during acquisition. The effect

<table>
<thead>
<tr>
<th></th>
<th>Problems 1 and 2</th>
<th>Problems 3 and 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A₁ A₂ A₃ A₄ A₅</td>
<td>A₁ A₂ A₃ A₄ A₅</td>
</tr>
<tr>
<td>A₁</td>
<td>25  75  93  90  84</td>
<td>69  65  77  81  73</td>
</tr>
<tr>
<td>A₂</td>
<td>17  21  65  77  45</td>
<td>35  50  77  84  60</td>
</tr>
<tr>
<td>A₃</td>
<td>11  17  35  74  34</td>
<td>23  30  72  37  23</td>
</tr>
<tr>
<td>A₄</td>
<td>10  11  23  26  18</td>
<td>19  16  28  28  23</td>
</tr>
<tr>
<td>A₅</td>
<td>10  11  23  26  18</td>
<td>19  16  28  28  23</td>
</tr>
</tbody>
</table>
Proportion of "wins"  

Fig. 5. Mean paired-comparison values for Experiments 6 and 7 (connected closed circles) together with predicted values (open circles).

of range of reward values during acquisition supports the impression gained from previous studies (e.g., Allen, 1972) that subjects do not master problems by learning relations between stimuli and numerical values in a paired-associate fashion.

B. EXPERIMENT 7. VARIATION IN MEAN REWARD DIFFERENCES WITHIN ACQUISITION PAIRS UNDER A PARTIAL INFORMATION CONDITION

The full information procedure of Experiment 6 leaves some doubt as to just what information the subjects were using from that displayed on each trial. Although a set of reward values was assigned to each alternative, some one member of the reward set being selected randomly on each trial when the subject chose the given alternative, nonetheless the full reward sets were displayed. It is possible that the subjects might, for example, have added up the
values in each set and then responded on the basis of these recoded values rather than on the basis of the sequence of rewards actually received. In order to clarify this point, in the present experiment we utilize the same procedures as those of Experiment 6 in all respects except that the full reward sets are never displayed. At the end of the trial on which, for example, the alternatives are A and B, the printout in front of the subject might take the form

A or B?
A
Reward = 5   Total = 61

Since the observation pairs of a given problem are presented at a rather rapid rate, and the subject sees on each trial only the single reward value that was selected by the computer from the appropriate reward set, it seems most unlikely that in a limited number of trials the subjects can possibly discover just what values belong to the various reward sets. Thus if the subjects are able under these circumstances to learn on the average to choose the higher paying alternative, they must do so by building up representations in memory of the relative positions of the stimulus alternatives on a reward scale. The task seems difficult, but not necessarily impossible since in earlier studies (Allen, 1972; Allen & Estes, 1972) it was found that even under simpler conditions when each alternative was associated with a single reward value the subjects did not always learn the exact values to the point of being able to verbalize them.

The stimuli and reward sets of the first two problems of Experiment 6 were used again, with a new group of eight subjects assigned to each. In view of the increased difficulty of the partial information procedure, 30 acquisition trials were given on each training pair before the administration of a five-trial test block on each of the test pairs.

In spite of the much increased difficulty of the task, rates of learning were not grossly different from those observed in Experiment 6. On Problem 1, with the relatively small differences in mean reward value within pairs, the subjects reached a terminal level of only 62% correct performance on the acquisition pairs by the end of the 30 training trials but on Problem 2, with the larger differences in reward values, learning was rapid and reached a terminal value of .92. On the critical first transfer trial, the subjects chose the higher valued alternatives 65% of the time on Problem 1 and 71% of the time on Problem 2.

To provide sufficiently stable data for quantitative comparison with the full information procedure, I have assembled in Table XIII
paired-comparison tables for each problem in terms of the percentage of choices of the row over the column alternative for the final acquisition block in the case of the training pairs and the full five transfer trials in the case of the test pairs. On the whole, both tables exhibit the typical orderly paired-comparison pattern. Also, if we consider the average values in the right-hand columns, it may be noted that for Problem 1 these are only slightly constricted as compared with those of Problem 1 in Experiment 6 and in the case of Problem 2 the quantitative agreement with the corresponding values for Problem 2 of Experiment 6 is striking indeed. It seems clear that the shift from full to partial information has made no qualitative difference in the mode of information processing.

C. EXPERIMENT 8. MEAN REWARD DIFFERENCES WITHIN DISJOINT TRAINING PAIRS

In both of Experiments 6 and 7, the training pairs were connected, in the sense discussed in relation to Experiment 5. Thus it remains to be determined whether the systematic transfer behavior depends on this property in the multiple-choice situation. The present experiment was designed to check on this question, and also to ascertain whether performance on pairs of former "winners" or former "losers" would vary as a function of mean reward differential within training pairs when differentials within test pairs were equated. The design is summarized in Table XIV.

Procedures were the same as those of Experiment 7. One group of eight subjects was assigned to Problems 1 and 3 and a second group to Problems 2 and 4, in each case receiving 30 acquisition trials on each training pair and 10 test trials. The test pairs were $A_1$ versus $A_3$,
TABLE XIV

DESIGN OF EXPERIMENT 8 IN TERMS OF MEAN REWARD VALUES PER TRAINING STIMULUS

<table>
<thead>
<tr>
<th>Training pairs</th>
<th>Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$A_1$ vs. $A_2$</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>6.3</td>
</tr>
<tr>
<td>$A_3$ vs. $A_4$</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>3.7</td>
</tr>
<tr>
<td>$A_5$ vs. $A_6$</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

$A_3$ versus $A_5$, and $A_1$ versus $A_5$ for the first group, and $A_2$ versus $A_4$, $A_4$ versus $A_6$, and $A_2$ versus $A_6$ for the second. The special feature of the design is that in Problems 1 and 3 the mean reward values of the high members (winners) of the training pairs were matched while the lower members differed, and in the other two problems, the lower members (losers) were matched while the higher members differed.

The principal result is that choices of the higher valued stimulus on transfer pairs prove to be substantially above chance for both pairs of winners (Problems 1 and 3) and pairs of losers (Problems 2 and 4), and to be directly related to differences in mean reward value between the members of a transfer pair: For Problems 1 and 3, mean reward differences of 2.7 and 5.4 yielded mean correct choice proportions of 76 and 90, respectively, whereas for Problems 2 and 4, differences of 1.3 and 2.6 yielded proportions of 60 and 66, respectively. As in previous experiments, more was learned in the case of former winners than former losers.

If we look, rather, at mean reward differences within training pairs, we find virtually no effect, overall correct choice proportions for Problems 1 and 3, respectively, being 81 and 80, for Problems 2 and 4, 63 and 60.

VI. Theoretical Interpretation of the Differential Reward Series

Taking the results of the present series of experiments together with others previously reported, we can assemble an array of rather
well-established facts that should sharply constrain the form of a model for acquisition and transfer in the multiple-choice situation.

1. In the case of two-choice problems, the rate of acquisition depends on the difference in reward value of the two alternatives when training is given with the partial information procedure but not with the full information procedure (Keller et al., 1965).

2. Rate of acquisition is largely independent of the ease of associating stimuli with specific reward values being, for example, quite comparable when there is a single reward value assigned to each stimulus and when there is a reward set with a probability distribution assigned to each stimulus.

3. Memory probes given during multiple choice acquisition show that specific reward values most often are not recallable by subjects until they are well into the final criterion series of 100% correct responding (Allen & Estes, 1972).

4. Probe tests calling for recall of reward values associated with individual stimuli yield orderly generalization gradients around the correct reward value following training under a partial information procedure (Estes, 1966).

5. When training has been given on only a subset of the pairs that can be formed from a set of stimuli, choice behavior at the end of training is comparable on training pairs and new transfer pairs, in both cases depending on the difference in reward value between the members of the test pair (Experiments 6 and 7).

6. More is learned concerning reward values associated with winning than with losing members of training pairs. However, transfer performance appropriately reflects differences in mean reward value within test pairs of former winners or former losers.

7. Choice performance on transfer trials depends strongly on the range of reward values prevailing during training but not on the differences in reward value between members of training pairs (Experiments 6–8).

The most general conclusion from the present series of studies is that transfer behavior following differential reward training has the character that would be expected if subjects had, in effect, learned to place stimuli on a scale of relative reward values. A central theoretical problem now is to formulate a model to account for the way in which the scale properties develop and the basis for them in the memory system. A number of hypotheses concerning both of these aspects of the problem that might be suggested by previous research and theory can be rather clearly refuted on the basis of the findings listed above.

First, in view of items 5 and 6, it is clear that subjects do not
simply learn to approach the higher and avoid the lower valued members of training pairs. Second, on the basis of items 1, 2, and 3, it seems equally clear that subjects do not in general associate specific numerical reward values with stimuli; and when they do so, at least in some cases this learning occurs too late to be importantly involved in the process of arriving at correct choice performance. Third, it does not appear that subjects master the problem by placing the stimulus alternatives on a preexisting scale of reward magnitudes; on such a hypothesis, we could account neither for item 1 nor for the fact that choice performance depends so strongly on the range of reward values involved in each specific problem of the series on which a given subject receives training (item 7).

This last observation implies that scale properties must be built up by some learning process that occurs independently on each problem of a series. What might be the nature of this process? One direction in which we might look for an answer is suggested by theoretical developments both in our previous analysis of probability learning and in the closely related area of verbal discrimination learning (Ekstrand, Wallace, & Underwood, 1966). Proceeding on analogy from those developments, one might speculate that, rather than rehearsing explicit reward values, subjects rehearse codes or labels for the stimulus alternatives, frequency of rehearsal being directly related to the placement of a stimulus in a range of possible values.

On this line of interpretation, the major question remaining is just how subjects achieve the result of rehearsing stimuli in proportion to their average reward values. A hypothesis considered by Medin (1972) requires the assumption of considerable long-term storage of reward information on the part of subjects. I shall proceed on a slightly different tack, allowing instead for a role of short-term memory.

Suppose we entertain the assumption that, following the appearance of the outcome on any trial, a representation in the form of an uncertainty distribution around the true reward value is entered in short-term memory, with some probability of being lost from memory (becoming unavailable) during each trial. In general, then, after the first few trials of an acquisition series, the subject will have available in short-term memory representations of the reward values associated with one or more stimuli. The critical assumption regarding rehearsal is that, once this state has been reached, when the subject observes the reward value given for the stimulus he chooses on a given trial, he compares this value with some one of the representa-
tions then active in short-term memory and rehearses a label for the chosen stimulus if its value is the larger. When more than one representation is in the active state, we assume that the one to be compared with the stimulus chosen on the given trial is sampled at random. The one exception is that, on a full information trial, the comparison is made between the displayed reward values rather than between representations in memory.

In general, the probability that the reward value of the stimulus chosen on a trial is greater than that of the one with whose representation it is compared will be an increasing function of the value of the given stimulus and consequently this modified hypothesis fits all of the same facts as the original. The advantages are two. First, we have a reasonably simple interpretation of the way in which the subject achieves the result of giving more rehearsal to stimuli with higher reward values. Second, since with a full information procedure the reward value of the stimulus not chosen from the pair presented on a trial is actually present in the display, it will always be the one compared to the chosen stimulus. Whenever the stimulus chosen has the higher value of the two it will, then, receive rehearsal regardless of the absolute value, so in the full information case, rate of acquisition is predicted to be independent of reward value, in accord with the results of Allen (1972) and Keller et al. (1965).

To obtain some more specific evidence regarding the merits of the present rehearsal hypothesis, I have conducted an analysis suggested by the analogy between the multiple choice and probability learning paradigms. In the case of each problem employed in Experiments 6 and 7, I consulted the listing of reward sets and determined for each stimulus the proportion of instances when a randomly selected member of its reward set would be larger in magnitude than that of a randomly selected member drawn from the other reward sets of the given problem. The value so obtained was assumed to represent the proportion of cases in which the given stimulus would be a “winner” in the hypothesized memory comparison engaged in by subjects during differential reward training. The proportion of cases in which the stimulus would be a “loser” was similarly determined. Then these two proportions for each stimulus were entered as estimates of $\alpha_i$ and $\beta_i$ in Eqs. (1)–(4), and the relative frequency model was applied to the data in the four paired-comparison matrices for Experiment 6 and the two for Experiment 7.

---

This statement applies only to the standard full information procedure in which only one reward value is assigned to each stimulus alternative.
The results of this theoretical analysis are summarized in Fig. 5 in the same form used for those of the probability learning experiments in Figs. 1-4. Once again, the agreement of observed and theoretical paired-comparison values appears quite promising. The estimates of the parameters \(a\) and \(b\) indicate some mixture of attentional strategies, but the generally high values of \(b\) (near unity for both problems of Experiment 7) suggest that in most instances subjects were encoding both wins and losses when engaged in memory comparisons. As in the case of probability learning, a direct approach to these data in terms of scaling theory would have encountered difficulties in that the observed choice behavior evidently again reflects a mixture of scales.

It is worth noting the perhaps counterintuitive result that the slopes of the functions in Fig. 5 are somewhat steeper for Problems 3 and 4 of Experiment 6, the problems with the more restricted ranges of reward values, than for Problems 1 and 2. In terms of the model, estimates of \(\phi\) are higher in the former case (.99 and 1.0 for Problems 3 and 4 versus .85 and .88 for Problems 1 and 2, respectively). However, it is a property of the model that differences in slopes of almost exactly the magnitude seen in the upper two panels of Fig. 5 would be predicted for the latter part of acquisition (\(\phi = .50\) to \(\phi = 1.0\)) even if learning rates (\(\phi\) values) were equal for all problems.

VII. General Discussion

For each of the two main types of tasks we have studied, probability learning and differential reward learning, we can find interpretations such that transfer behavior conforms to expectations on the assumption that the subject has placed the choice alternatives on a psychological scale and chooses from test pairs on the basis of scale values. The functional properties of the system yielding these scale properties provides economy of information storage and precision in transfer performance.

But the question remains whether the conception of a psychological scale should be regarded as a fundamental one for theoretical purposes or whether it is derivative to more basic theoretical ideas. We have seen that, in the case of probability learning, the scale does not agree with the traditional conception of probability in that it is not always monotonely related to objective probability. Further, in cases where the function is monotone, it does not generally prove to
be linear. It is apparent from Eq. (8) that observed choice probability should be curvilinearly related to $v_k$. However, we can transform the equation into the form

$$\frac{1 - P_{ij}(i)}{P_{ij}(t)} = \frac{P_{ij}(j)}{P_{ij}(i)} = \frac{v_j}{v_i},$$

and, taking logarithms of both sides,

$$\log P_{ij}(j) - \log P_{ij}(i) = \log v_j - \log v_i. \quad (9)$$

The quantity on the left can be directly estimated from the observed choice proportions, and the average value for alternative $A_i$ should yield a linear function of $\log v_i$. With $v_i$ taken to be $\omega_i^2$ for probability learning and $\beta_i/\alpha_i$ for multiple-choice learning, these plots (not reproduced here) prove to be linear in neither instance. In several cases, for both types of data, the functions appear to be better described by two straight lines of different slopes, possibly reflecting a mixture of two scales as suggested by our analysis in terms of the learning model. It would seem sensible to defer further pursuit of this analysis until experiments have been run under conditions which either yield homogeneity of attentional strategies for groups of subjects or provide adequate data for theoretical analysis of the choice behavior of individual subjects.

On the whole, the evidence from the present studies indicates that the psychological scales manifest in transfer performance do not represent preexisting structures that subjects learn to use in a given situation. Rather they appear to represent a type of organization of information in memory that takes form in the course of learning in a given task situation.

Consideration of the effects of various experimental manipulations leads to the conclusion that the psychological scales represent memory structures built up by differential rehearsal of stimulus codes. These structures have some of the properties of the frequency count assumed by Ekstrand et al. (1966) in their frequency theory of verbal discrimination learning, and come still closer to the notion of a relative frequency indicator (Estes, 1976) reflecting memory for relative frequencies of event categories. In the case of probability learning, the nature of the scale structure closely reflects the relative frequencies of outcomes to which subjects selectively attend. In the case of multiple-choice learning, the structure may well reflect the
relative reward values that subjects come to expect as a consequence of their learning experience, but not necessarily the different amounts of reward received for alternative choices on past occasions. In both cases, close attention to conditions of selective attention, event coding in memory, and retrieval from long-term memory are necessary in order to specify the type of scale that will take form.

REFERENCES


