# Probability Matching: Encouraging Optimal Responding in Humans

EDMUND FANTINO and ALI ESFANDIARI
*University of California, San Diego*

**Abstract** Two hundred undergraduate students participated in a repeated-trials binary choice procedure in which choice of one outcome was correct on 75% of trials. Subjects received 192 trials and were divided into five conditions: (1) control; (2) subjects were given the actual probabilities; (3) subjects were told if they did well they could leave early; (4) competition condition; (5) midway through the task subjects were asked to recommend a strategy for another subject. Half of the subjects in each group were told that the best they could do was to be correct on 75% of the trials. This manipulation permitted assessment of the hypothesis that subjects in probability-matching tasks are seeking a strategy that will be correct on 100% of the trials. The results partially confirmed this hypothesis. In addition, two of the variables improved performance significantly (giving probabilities and asking subjects to recommend a strategy). However, while subjects in all groups improved significantly over trials, optimal choice did not occur in this task.

**Résumé** Deux cents étudiants du premier cycle universitaire ont participé à des essais répétés de procédure de choix binaire, pour lesquels l'une des réponses était exacte dans 75 % des essais. Les sujets étaient soumis à 192 essais, qui se divisaient en cinq conditions expérimentales : 1) contrôle; 2) on donnait aux sujets les probabilités réelles; 3) on disait aux sujets que s'ils réussissaient bien, ils pourraient partir avant la fin des essais; 4) compétition; 5) à mi-chemin entre le début et la fin de l'exécution de la tâche, on demandait aux sujets de suggérer à un autre sujet une stratégie. La moitié des sujets de chaque groupe étaient informés que le mieux qu'ils pourraient faire serait de choisir la bonne réponse dans 75 % des essais. Cette manipulation permettait d'évaluer l'hypothèse selon laquelle les sujets qui exécutent des tâches de correspondance en fonction de la probabilité cherchent à adopter une stratégie qui leur permettra de répondre correctement à 100 % des essais. Nos résultats confirment en partie cette hypothèse. De plus, deux des variables ont permis d'améliorer la performance de façon significative (« donner les probabilités de réussite » et « demander aux sujets de suggérer une stratégie »). Néanmoins, même si les sujets de tous les groupes amélioraient considérablement leur performance à mesure qu'ils participaient aux essais, le taux optimal de choix de réponse n'a pas été observé au cours de cette tâche.

When human subjects are presented with repeated identical binary choices, each associated with a constant payoff likelihood, they tend to match their choices to the arranged probabilities (e.g., Humphreys, 1939) instead of maximizing their payoffs by always choosing the stimulus with the higher reinforcement likelihood. Thus, if one event occurs 75% of the time, subjects tend to choose that outcome on 75-80% of the trials rather than choosing it on every trial. This phenomenon is termed "probability matching" and is typically viewed as an error on the subject's part. Each time the subject predicts the less likely payoff, his or her probability of being correct (and earning whatever reward comes from being correct) declines, thereby reducing the overall expected payoff. Probability matching is an extremely robust phenomenon (e.g., Myers, 1976; Shanks, 1990), persisting over hundreds of trials. While Goodie and Fantino (1999) found a gradual transition towards optimal responding over 1,600 trials, as Fantino has noted "Life rarely offers 1,600 trials" (Fantino, 1998, p. 213). Other than conducting over 1,000 trials, there have been procedures that have encouraged somewhat more optimal responding, including sufficiently directive instructions (e.g., in the most extreme case, Braveman & Fischer, 1968, told subjects that the task was entirely random, that one option was correct more often than the other, that optimal behaviour was to ascertain which option was correct more often and choose this option exclusively, and that this was the subject's task). Fantino (1998) reported pilot data (collected with Greg Zarow) implicating several variables that may affect optimal responding. Some of these variables are assessed in the present experiment. That the use of strategies often interferes with optimal decision-making by humans is a point stressed by several investigators. For example, both Arkes and Ayton (1999) and Goodie and Fantino (1996) have stressed that nonhumans often do better than humans in comparable reasoning tasks because they lack the potentially interfering history that human subjects bring to the task. Thus, Hartl and Fantino's (1996) pigeons selected much more optimally in a study investigating the utilization of base rates than did Goodie and Fantino's (1995, 1996) college students in a procedurally equivalent task.

One possibility that has not been evaluated involves the likelihood that subjects in a probability-matching task are attempting to devise a strategy in which they are correct on close to 100% of the trials. Given subjects' rich history of problem-solving this is a reasonable assumption (e.g., Levine, 1974). Indeed, Weisberg (1980), in his discussion of problem solving, notes: "So long as the subject believes that some correct sequence will produce *errorless* performance, the fact that one stimulus is reinforced more than the other is irrelevant and has no effect on subsequent choices" (p. 304). A study by Yellott (1969) supports this notion. Subjects in his binary choice experiment had to predict which of two lights would appear. For many trials the more frequent light appeared on 80% of the trials. Subjects matched their predictions to this frequency. Then, on subsequent trials, without the subjects' knowledge, the light appeared wherever the subject predicted it would be, guaranteeing a 100% success rate. Subjects continued to select the previously more frequent light on 80% of the trials. Interestingly, the subjects' description of their decision strategies revealed complex sequences of choices, supporting the likelihood that subjects had been seeking causal sequences throughout the experiment and were misled into thinking that they had finally succeeded. Yellott compared the complex strategies to the superstitious behaviour found in the operant laboratory (Yellott, 1969, p. 579; Catania & Cutts, 1963, also reported evidence for inappropriately complex strategies in a study of operant choice). Wolford, Miller, and Gazzaniga (2000) have findings in split-brain patients implicating the left hemisphere's role in hypothesis formation, pointing to the possibility that the left hemisphere of humans houses a cognitive mechanism that interprets past occurrences. They note:

> So why do humans choose a less optimal strategy than rats? Our view is that humans believe there is a pattern, even if told the sequence is random, and they attempt to figure out the pattern. Any reasonable pattern hypothesized by the subjects would have to match frequency if it were to be a correct hypothesis. Perhaps animals other than humans adopt a more optimal strategy than humans in this paradigm, because they are not as hindered by the tendency to search for and posit causal hypotheses, (p. 2).

This conclusion is consistent with those of Arkes and Ayton (1999), Goodie and Fantino (1996), Levine (1974), and of Weisberg (1980) cited above. Perhaps then by simply informing subjects that the best possible performance is being correct on x% of trials (where x equals the percentage corresponding to the more likely payoff), subjects will optimize. This possibility may be enhanced by also informing subjects of the probabilities that each of the two outcomes will be correct. The present study assessed this account of probability matching. At the same time it explored additional variables that might affect degree of probability matching, especially as a function of being told the optimal score. Thus we also manipulated subjects' motivation (by making the task a competitive one or by giving them the opportunity to leave the experiment early if they did well) and encouraged some subjects to reflect on the task by asking them to recommend a strategy for others.

## Method

### PARTICIPANTS

Two hundred (82 male, 118 female) young adults (mean age = 20.4, SD = 2 years) served in this study. They were students enrolled in psychology and cognitive science courses. Participation was rewarded with course credit.

### APPARATUS

Sessions were conducted in a small testing room with ambient lighting. Participants sat at small computer desks facing a PC display with a keyboard that had two active keys. Visual stimuli, in the form of blue and green circles, were presented on the computer screen. Each colour was associated with a probability of reinforcement. The two active keys (s and l) had separate functions, with each representing a unique left or right stimulus choice. Once a choice had been made, the display changed in a more significant fashion, indicating the outcome of the choice. The computer processed all data recording.

### PROCEDURE

A binary-choice probability learning procedure was utilized. Participants were assigned randomly to one of five conditions, each consisting of 40 subjects in a

TABLE 1
Mean Choice Proportion (and Standard Error) for the Rich Choice

| | Control | Probabilities | Competition | Leave Early | Strategy |
|---|---|---|---|---|---|
| No Instructions | | | | | |
| (N = 20 per condition) | | | | | |
| First 96 Trials | .75 (.03) | .82 (.03) | .79 (.03) | .74 (.02) | .80 (.02) |
| Last 96 Trials | .85 (.03) | .87 (.03) | .86 (.03) | .84 (.03) | .91 (.03) |
| All Trials | .80 (.03) | .85 (.03) | .82 (.03) | .79 (.02) | .85 (.02) |
| Instructions | | | | | |
| (N = 20 per condition) | | | | | |
| First 96 Trials | .76 (.02) | .88 (.02) | .77 (.03) | .80 (.02) | .79 (.02) |
| Last 96 Trials | .85 (.02) | .93 (.02) | .84 (.03) | .91 (.02) | .87 (.02) |
| All Trials | .80 (.02) | .90 (.02) | .80 (.03) | .85 (.02) | .83 (.02) |
| Collapsed Across Instructions | | | | | |
| (N = 40 per condition) | | | | | |
| First 96 Trials | .75 (.02) | .85 (.02) | .78 (.02) | .77 (.02) | .79 (.01) |
| Last 96 Trials | .85 (.02) | .90 (.02) | .85 (.02) | .87 (.02) | .89 (.02) |
| All Trials | .80 (.02) | .87 (.02) | .81 (.02) | .82 (.01) | .84 (.01) |

between-subjects design. Each person was also assigned randomly to either an Instructions or a No Instructions treatment within each of the five conditions, producing 10 subgroups of 20 subjects each. Each session consisted of two 96-trial blocks (for a total of 192 trials) of a binary choice between the two coloured circles. For all subjects, a rest period of at least 60 seconds was interposed between the two 96-trial blocks. After 60 seconds, the program waited until subjects were ready to continue. The probabilities used were 75% and 25%, and remained constant across conditions. The stimulus and side associated with the higher probability was counterbalanced across participants. Reinforcement consisted of a "point," and the computer display indicated clearly the participant's cumulative total. The inter-trial interval was set at five seconds, at which time the computer display was darkened.

In the control condition, participants read written instructions (see below) to earn as many points as possible. The second condition was identical to the control condition, with the only exception being that the participants were told the actual probabilities and the side and colour with which the probabilities were associated. In a third condition, two participants were simultaneously given instructions that they were in direct competition, with the task being to earn more points than earned by the other person. In the fourth condition, subjects were instructed that the length of the study was one hour (it actually tended to last about 30 minutes), and that they could leave early if their score fell in the top 5% of participants studied up to that point. The fifth condition was identical to the control condition, with the exception that participants were instructed to recommend a strategy to another participant during the rest break between the two 96-trial blocks.

In addition to the differences between the five main conditions, half of the participants in each of the five conditions were instructed that obtaining 75% of the available points would be a perfect score. The aim here was to minimize the likelihood that subjects' probability matching reflected an attempt to find a "solution" that enabled them to be correct on closer to 100% of the trials.

The written instructions included the following for all conditions, and were the sole instructions present for the control group.

> Two coloured circles will be presented on the computer screen. You can choose freely between the two, pressing the "S" key to select the left circle and the "L" key to select the right. On each presentation of these two circles, one will lead to a point. Your task is simply to earn as many points as you can.

For the "Probabilities" condition, a variation (according to the counterbalancing which was taking place) of the following was also included:

> Responding to the blue circle on the left will provide a point 75% of the time, while responding to the green circle on the right will provide a point 25% of the time.

For the "Competition" condition, in which pairs of participants were verbally instructed that they were going to be competing on a computer task, the written instructions also included:

> You will be competing with another subject running at the same time in a different room.

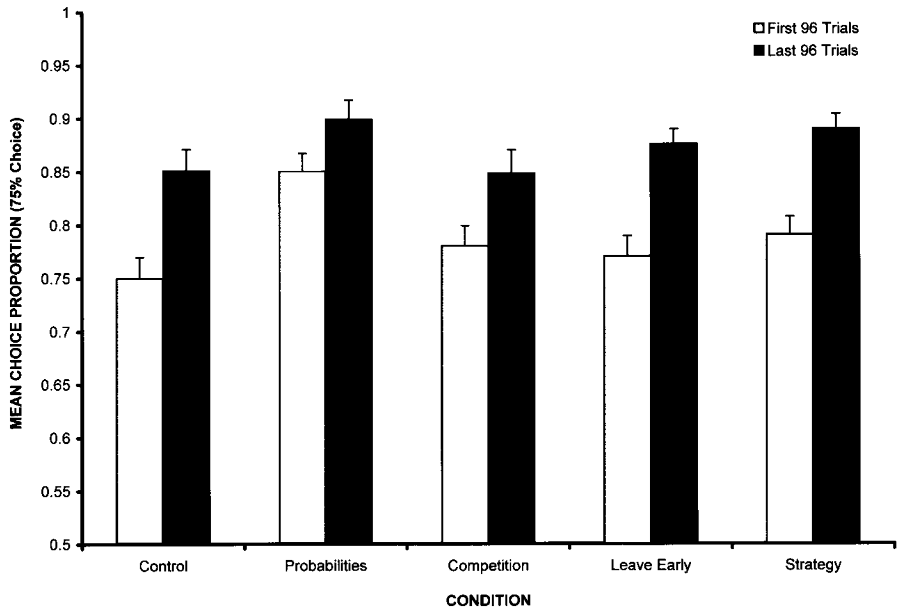


*Figure 1.* For each of five groups, the mean choice proportion expressed in terms of the outcome with the more likely payoff ("75% choice" on ordinate), for both the first and second ("last") block of 96 trials. Standard errors of the mean are shown.

For the "Leave Early" condition, the subjects were verbally reminded that the study was one hour (the sign-up sheet which recruited them also indicated the study was one-hour long) and the following was added to the written instructions:

> If your score falls within the top 5%, you will get to leave early, skipping the second half of the experiment. There will be multiple rest periods. The computer will automatically determine if your score meets the 5% requirement during the second rest period.

For the "Strategy" condition, the following was added to the written instructions:

> At some point during the experiment, the program will pause and ask you to get the experimenter. The experimenter will then ask you to recommend a strategy to another participant.

Half of the subjects in each condition, including the control, also received the following written instructions:

> Obtaining 75% of the available points is considered a perfect score.

## Results

The participants' choice proportions are shown in Table 1 and Figure 1. Results in the table give the means and standard errors of the means for the first 96 trials, for the last 96 trials, and for all trials, in each condition. For all statistical analyses, the $p$-value was .05. Although the instructions that the subjects could do no better than be correct on 75% of the trials appeared to enhance selection of the more likely outcome in some conditions (especially "probability" and "strategy"), this enhancement was not statistically significant according to an analysis of variance, $F(1, 199) = 1.07$, n.s. Thus, the means in Figure 1 are collapsed over the "Instructions" and "No Instructions" treatments. The figure shows the mean choice proportion of the more likely outcome (the "75% Choice" indicated on the ordinate) averaged over the 96 trials in each of the two blocs of 96 trials for each of the five conditions studied.

As is evident from the figure, subjects improved over trials. A repeated-measure ANOVA indicated that this improvement was significant across the two blocs of trials, $F(1, 199) = 132.85$. An ANOVA showed that there was also a significant effect of conditions, $F(4,196) = 3.02$. A Neuman-Keuls post-hoc test revealed that the Probability condition (mean choice of more likely outcome over trials = .87) was significantly greater than each of the other conditions, and that the Probability and Strategy (.84) conditions were each significantly higher than the Control condition (.80).

Although only two individuals performed at an optimal level (choice of more likely outcome on all trials) during the first bloc of 96 trials, 23 of the 200 subjects performed at an optimal level during the second bloc of trials. Of these almost half (11) came from the "Probability" group and only one from the Control group (five were from the "Strategy" group, four from the "Competition" group, and two from the "Leave-Early" group). A chi-square test indicated that this distribution differed from chance ($C^2$ = 15.03, 4 df). Ryan's procedure showed that there were significant differences between the Probability group and the Control group and between the Probability group and the Leave-Early group. Of the 23 subjects choosing optimally in the second bloc, 9 came from "No-Instructions" groups and 14 from "Instruction" groups. All subjects in the Strategy group indicated awareness that one side delivered more points. However, despite this insight, most still recommended strategies that included occasionally sampling the 25% side (e.g., "Stay on the left side, since you get more points, but once in a while check the other side").

## Discussion

The present results are consistent with many others in showing that probability matching, a nonoptimal decision strategy, rather than maximization, the optimal strategy, is the typical response of educated subjects to a repeated binary decision task. While subjects increasingly chose the more likely outcome across 192 trials, and therefore were no longer probability matching, their mean choice proportion for the more likely payoff did not reach 100% in any condition. In this procedure cue similarity was maximal in that subjects were presented with the identical pair of stimuli regardless of which was correct on a given trial. According to Massaro's (1969) model for discrimination learning, such highly confusable trials should make it difficult for testing of strategies, resulting in matching. We reasoned that, by informing subjects about the maximal payoff, we might minimize one source of confusion about the task, thereby encouraging more optimal responding. Specifically, our hypothesis had been that by informing subjects that they could be correct on no more than 75% of trials, subjects would not seek a solution that led to a 100% rate of success, a plausible account for past findings of probability matching, and one consistent with the research of Levine on more complex problems (e.g., Levine, 1974) and with a hypothesis proposed by Weisberg (1980) and cited in the Introduction. According to this view, informed subjects would be more likely to select the more likely outcome on all trials (and achieve the maximal success rate of 75%). Instead, optimal behaviour did not occur whether or not subjects were given these instructions. Perhaps subjects did not believe our instructions. In any event, even with the information we provided, they did not respond optimally.

We undertook four additional manipulations that might have improved subjects' performance in the direction of becoming more optimal (two that should have increased motivation, a third providing the actual probabilities, and a fourth asking them to advise another subject). Although all had at least a small effect in the expected direction, only two produced performance that was statistically superior to that of the control group (the probability and strategy conditions). The performance most approaching optimal was that of the Probability group: Subjects told the actual probabilities chose the more likely outcome on 88% of trials. Moreover, if we look at only the second block of 96 trials, subjects in the Probability group chose the more likely outcome on 90% of trials (with 11 subjects performing at 100%). Finally, the 20 subjects who were given the probabilities and the instruction that 75% was the optimal score, chose the more likely outcome on 93% of the trials during the second block of 96 trials (with 6 of these 20 subjects at 100%). Thus, those individuals given a veridical statement of the probabilities and of the optimal success rate perform at a level intermediate to matching and maximization initially (88% on first 96 trials) and soon begin to approach maximization (93% on the second 96 trials). In other words an approximation to optimal behaviour may be achieved in fewer than 200 trials without actually telling subjects that they should choose the more likely outcome on each trial. The other treatment that enhanced performance involved asking subjects to advise others. Presumably this had the effect of encouraging them to reflect on the contingencies; once they articulated a more effective strategy for another to follow, they were somewhat more likely to follow that same strategy themselves.

It is unclear whether additional improvement would have occurred had the task been extended beyond 192 trials. Wolford (personal communication) studied 500 trials and reported little change after the first 100 trials. On the other hand, in a related task involving base-rate neglect, Goodie and Fantino (1999) reported a gradual transition towards optimal responding over 1,600 trials.

In summary, despite the success of two of the independent variables studied here in affecting decision-making in a repeated-trials binary choice (probability and strategy conditions), and despite the improvement shown by all subjects across trials, the most compelling conclusion warranted by the present data is that subjects' decision making in this task is not optimal. However, subjects informed of the optimal success rate and of the true probabilities began to approach optimal behaviour after about 100 trials, giving limited support to the view that probability matching tends to occur because subjects seek a strategy that will succeed on

100% of the trials.

The authors thank Dr. Stephanie Stolarz-Fantino for constructive comments on earlier drafts of this article. Research and preparation of this article was supported by National Institute of Mental Health Grant MH57127 and National Science Foundation Grant IBN-9870900.

Correspondence concerning this article should be addressed to Edmund Fantino, Department of Psychology-0109, University of California, San Diego, La Jolla, CA 92093-0109. (E-mail: efantino@ucsd.edu).

**References**

Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower human animals? *Psychological Bulletin, 125*, 591-600

Braveman, N. S., & Fischer, G. J. (1968). Instructionally induced strategy and sequential information in probability learning. *Journal of Experimental Psychology, 76*, 674-676.

Catania, A.C., & Cutts, D. (1963). Experimental control of superstitious responding in humans. *Journal of the Experimental Analysis of Behavior, 6,* 203-208.

Fantino, E. (1998). Judgment and decision making: Behavioural approaches. *The Behavior Analyst, 21,* 203-218.

Goodie, A. S., & Fantino, E. (1995). An experientially derived base-rate error in humans. *Psychological Science, 6,* 101-106.

Goodie, A. S., & Fantino, E. (1996). Learning to commit or avoid the base-rate error. *Nature, 380*, 247-249.

Goodie, A. S., & Fantino, E. (1999). What does and does not alleviate base-rate neglect under direct experience. *Journal of Behavioral Decision Making, 12*, 307-335.

Hartl, J. A., & Fantino, E. (1996). Choice as a function of reinforcement ratios in delayed matching to sample. *Journal of the Experimental Analysis of Behavior, 66*, 11-27.

Humphreys, L. G. (1939). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology, 25*, 294-301.

Levine, M. A. (1974). A transfer hypothesis, whereby learning-to-learn, Einstellung, the PREE, reversal-nonreversal shifts, and other curiosities are elucidated. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium*, Potomac, MD, 289-305. Hillsdale, NJ: Erlbaum.

Massaro, D. W. (1969). A three state Markov model for discrimination learning. *Journal of Mathematical Psychology, 6*, 62-80.

Myers, J. L. (1976). Probability learning and sequence learning. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes,* 171-205. Hillsdale, NJ: Erlbaum.

Shanks, D. R. (1990). Connectionism and the learning of probabilistic concepts. *The Quarterly Journal of Experimental Psychology, 42A*, 209-237.

Weisberg, R. W. (1980). *Memory, thought, and behavior*. New York: Oxford University Press.

Wolford, G., Miller, M. B., & Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. *The Journal of Neuroscience, 20,* RC64, 1-4.

Yellott, J. I., Jr. (1969). Probability learning with noncontingent success. *Journal of Mathematical Psychology, 6*, 541-575.