PROBABILITY AND DELAY IN COMMITMENT

HOWARD RACHLIN, ANTHONY CASTROGIOVANNI, AND DAVID CROSS

STATE UNIVERSITY OF NEW YORK AT STONY BROOK

In the first stage of a two-stage choice, human subjects chose between probabilistic access to a second choice (between a small high-probability reward and a large low-probability reward) and commitment to the large low-probability reward. When confronted with the second-stage choice, subjects strongly preferred the small high-probability reward. When the first-stage probability (of access to the second stage) was high, subjects strongly preferred the path leading to the choice in the second stage. But when the first-stage probability was low, subjects committed themselves to the large low-probability reward. These results parallel those obtained by Rachlin and Green (1972) with pigeons and constitute some evidence that probabilities may be interpreted as delays.

Key words: choice, reinforcement delay, humans

Rotter (1954) has suggested that probability and delay of reward are subjectively equivalent. He said that people represent promised rewards subjectively as probabilities; the sooner the reward is promised, the higher the probability of its subjective representation. According to Rotter, the reason people with poor self-control consistently choose small-immediate over large-delayed rewards is because the probabilities of the larger (delayed) rewards are subjectively low. In other words, people with poor self-control believe that promised delayed rewards may never be delivered. Rotter did not extend his theory to behavior of nonhuman animals, but presumably a low subjective probability of reinforcement could result from one that has been delayed in the past (is part of the animal's reinforcement history) as well as from one that is verbally promised in the future.

Mischel and Grusec (1967) found that human children were equally good or equally poor in choosing large over small rewards when the large rewards were relatively delayed as when they were relatively improbable. Stevenson (1986), observing people's business decisions, Navarick (1987), studying people's choices between reinforcement schedules, and Rachlin, Logue, Gibbon, and Frankel (1986), manipulating the intertrial interval of a human gambling task, also found equivalencies

between delay and probability of reward. The objectives of the present experiment are first to extend the delay-probability equivalency to a commitment paradigm and second to argue, contrary to Rotter, that the fundamental variable underlying the common effect is delay, not probability.

Figure 1a diagrams the paradigm used by Rachlin and Green (1972) with pigeons as subjects. The present study with human subjects, diagrammed in Figure 1b, replaces the delays of the previous study with probabilities—long delays with low probabilities, short delays with high probabilities. The pigeons, at point A of Figure 1a, strongly preferred the small-immediate reinforcer ($a = 2$ s, $d = 0$ s) to the large-delayed reinforcer ($a = 4$ s, $d = 4$ s). When t , the time between Choice B and Choice A, was brief (less than 4 s), pigeons preferred the path leading to Choice A and subsequently chose the small-immediate reinforcer. (The pigeons chose among the alternatives by pecking appropriate keys.) However, when t was relatively long (more than about 4 s), the pigeons preferred the lower path, leading to Point C at which only one alternative was available—the large-delayed reinforcer. By choosing the lower path the pigeons committed themselves to the larger, albeit more delayed, reinforcer.

If humans' choices among probable rewards correspond to pigeons' choices among delayed reinforcers, the subjects of the present experiment, at A in Figure 1b, would prefer the small, highly probable, reward ($a = \$1$, $p = 17/18$) to the larger, less probable, reward. If

This research was supported by a grant from the National Science Foundation. Send reprint requests to Howard Rachlin, Psychology Department, State University of New York, Stony Brook, New York 11794-2500.

Fig. 1. Part (a) shows a paradigm of the choice structure for a delay-commitment paradigm (Rachlin & Green, 1972), Part (b) shows the analogous paradigm for the present experiment with human subjects choosing among monetary rewards of various probabilities.

q were a high probability (corresponding to a short delay) the subjects, at B, would prefer the path leading to A, thence to choose the small, highly probable, reward. However, if q were a low probability (corresponding to a long delay) the subjects, at B, would prefer the path leading to C over that leading to A and, like the pigeons, commit themselves to the larger reward.

METHOD

Subjects

Eleven Stony Brook undergraduates, 9 women and 2 men, volunteered as subjects. The experiment was run individually for each subject.

Apparatus

The subject stood in front of a spinner with 18 white sectors labeled 1 through 18. The spinner provided a way to unpredictably select one of the 18 numbers. The experimenter sat across the table from the subject. Between the

Fig. 2. Arrangement of cards and spinner. The numbers in parentheses correspond to numbers on the sections of the 18-section spinner.

subject and the spinner were five cards arranged as in Figure 2, a lower rank of two cards, X and Y, and an upper rank of dollar cards. The cards were labeled as in the figure. The numbers in parentheses on the cards refer to the numbers on the spinner.

Procedure

At the beginning of the experiment each subject was given 20 poker chips, 10 red and 10 blue, and the experimenter read the instructions (included below). Each choice was indicated by putting a chip on a card. In the first stage the subject could put a red or a blue chip on Card X or Card Y. Then the experimenter spun the spinner. If the spinner did not land within the range of numbers on the card that corresponded to the color of the chip played, the experimenter took the chip and a new trial began immediately with another chip. If the spinner did land within the winning range, the trial proceeded to the next stage. If, in the first stage, the subject had chosen X and won, the subject could choose either the \$1 or the \$4 bet by moving the chip from Card X to either of the dollar cards above it. If in the first stage, the subject had chosen Y and won, the chip could be moved from Card Y to the \$5 card above it. Then the experimenter spun the spinner again. After a win on one of the dollar cards, the experimenter took the red or blue chip and gave the subject a number of white chips equal to the dollar amount won. After a loss, the experimenter simply took the chip. Thus each trial consisted of one or two spins corresponding to the first and second

choice stages. During the first stage the chip was on Card X or Card Y. A win at the first stage led to the second stage with the chip on the \$1, \$4, or \$5 card. A win at the second stage consisted of a monetary reward. The session consisted of 20 trials, one for each red or blue chip. At the end of the experiment, the subject cashed in the white chips for an equivalent amount of dollars.

These are the instructions read to the subjects:

You have in front of you twenty tokens, ten red and ten blue. Each token represents a chance to win money that you can keep at the end of this experiment. You will have twenty chances, or trials, to win,—one trial for each token. When your tokens are used up, the session is over.

Each trial will have two stages:

In stage one, I will ask you to put a token on one of these cards. You may choose to use a blue or red token and you may put either token on either card [E points to X and Y]. After you have placed your token on a card, I will spin the spinner. If you have bet a red token and the spinner lands on numbers 1–15, you will be able to go on to the next stage. However, if the spinner lands on 16, 17, or 18 you will lose the token and that trial will be over. If you have bet a blue token and the spinner lands on numbers 1, 2, or 3, you will be able to go on to the next stage. If, however, the spinner lands on anything other than numbers 1–3, the trial will be over and you will lose the token.

Stage two commences only after you have successfully completed stage one. If your token (red or blue) was on X in stage one, you now have another choice to make: If you now move your token to this card [E points to \$1 card], I will spin the spinner and if it lands on numbers 1–17 you will win a white token that is worth \$1 to you; however, if instead of moving your token to the \$1 card, you had placed it here [E points to the \$4 card], if the spinner stops on numbers 1–5 you will win four white tokens (worth \$4). In either case, if the spinner lands on any other number than those which are on the card, you will lose your token and the trial will be over.

If your token (red or blue) was on Y in stage one, you must now move it to this card [E points to \$5 card]. The trial will be essentially the same as with the X card, except that the spinner must now land on numbers 1–4 for you to win and your winnings will be five white chips (worth \$5 to you). Again, if the spinner lands on anything other than numbers 1–4, you will lose your token and the trial will be over.

At the end of the session, I will give you the number of real dollars that your white chips represent. This money is yours to keep.

You may play the red or blue chips in any sequence and on either card (X or Y) at your discretion. There are no rules as to where you must place your token or which color it must be.

Please take your time as we play and consider every trial carefully, but do not try to make mathematical calculations. However, we do want your considered judgment on each trial.

Are there any questions?

Let us try a few trials for practice before beginning [E runs through each possibility once, i.e., red and blue on X, red and blue on Y, and each of the three money cards].

The probability q in Figure 1b is that corresponding to the red or blue chip. The odds of proceeding from Choice B to Choice A or Point C were 15/18 when the subject played a red chip and 3/18 when the subject played a blue chip. Because the subject had 10 red and 10 blue chips, the experiment consisted of 10 trials with $q = 3/18$ and 10 trials with $q = 15/18$. This method of having the subjects themselves select trial order was chosen because pilot experiments of ours as well as published accounts of human laboratory analogs to animal experiments (e.g., Hayes, Brownstein, Haas, & Greenway, 1986) indicate that corresponding results are more likely when people's tasks are made more complicated and varied than corresponding animal tasks. For similar reasons, the two larger rewards (\$5 and \$4) were not identical (although their expected values were identical).

RESULTS

The median amount won by the 11 subjects was \$7 for a session that lasted a median time of 5 min (not counting time spent on instructions). Subjects tended to bet most of their red chips (the high-probability, more valuable, ones) on the early trials and most of their blue chips on the later trials. More than twice as many red chips as blue chips were bet on the first 10 trials (74:36) whereas, necessarily, the reverse applied for the second 10 trials. Three subjects bet all of their red chips followed by all of their blue chips, but most varied their bets. Most subjects saved some red chips for

the last few trials. On the very last trial 7 subjects had a red chip and 4 had a blue chip left to bet.

Because of the small sample size, the actual obtained payoff frequencies only approximated the nominal odds as represented on the spinner. For instance, the nominal odds of passing to the second stage on red-chip bets were 15/18 or 0.83. The actual frequencies for the 11 subjects were 0.6 for 1 subject, 0.8 for 3 subjects, 0.9 for 4 subjects, and 1.0 for 3 subjects, an average of 0.87 or about 16/18. Nevertheless, there is evidence that in the short run, as in the brief 5-min sessions of this experiment, human behavior is controlled by instructions (nominal odds) rather than by experienced contingencies (Hayes *et al.*, 1986). Furthermore, instructions are likely to be more dominant when, as in the present case, they are not contradicted by experienced frequencies.

At Point A, 8 of the 11 subjects preferred the small high-probability reward ($a = \$1$, $p = 17/18$) to the large low-probability reward ($a = \$4$, $p = 5/18$), 1 subject was indifferent, and 2 subjects preferred the large reward. The average fraction of choices of the small high-probability reward was 0.72, significantly different from .5 at the .005 level ($t = 3.28$, $df = 10$, one-tailed test). At Point B 10 of the 11 subjects placed more of their red chips than blue chips on Card X. (The other subject placed an equal number of chips on both cards.) An average of 8.18 of the 10 red chips ($q = 15/18$) and an average of 4.73 of the 10 blue chips ($q = 3/18$) were placed on Card X (leading to Choice A). This difference is significant at the .001 level ($t = 5.42$, $df = 10$, two-tailed test). Thus, a very strong preference for small high-probability rewards was changed to indifference (preference was slightly but not significantly reversed) when a prior low-probability commitment alternative was introduced. Figure 3 shows the choice of the larger reward at Point A and the choice of commitment at Point B for the two values of q studied.

Because subjects tended to bet their red chips first, it is conceivable that the increase in commitment from $q = 15/18$ (red chips) to $q = 3/18$ (blue chips) was an effect of order or time—subjects might have learned commitment over the course of the experiment. This possibility may be tested by examining the last

four trials alone. Because many subjects saved one or two red chips to use at the end of the experiment, the number of red and blue chips bet during the last four trials by the 11 subjects was equal at 22 each. Of the 22 red chips, 18 (82%) were bet on the noncommitment alternative (Alternative X). Of the 22 blue chips, 9 (41%) were bet on the noncommitment alternative. The corresponding percentages for all 20 trials are 81% and 47%.

During the last four trials, a total of 15 choices was made between the \$1 and \$4 alternatives. Of those, 11 (73%) were of the small low-probability alternative, compared to 72% for the experiment as a whole. Thus, behavior during the last four trials reflected behavior during all 20 trials of the experiment. Because, over those four trials, the order tended to be blue-red (on the 17th trial, 7 of the 11 subjects bet a blue chip; on the 20th trial, 7 of the 11 subjects bet a red chip) the overall results are unlikely to have been due to the overall red-blue order.

DISCUSSION

These results parallel those obtained with pigeons choosing among rewards of various amounts and delays. Increases of t in the pigeon experiment and decreases of q in the present experiment had the same effect—more commitment. Although some individual subjects came to make a strong preponderance of commitment responses, the average proportion of such responses (here, 53%) was only slightly and nonsignificantly above chance level. However, both humans and pigeons significantly increased their overall proportion of commitment responses as q was lowered (t raised) and some subjects in both experiments made a large number of commitment responses when q was low (t was high). One difference between the present results and those with pigeons is that, at A, the pigeons pecked on more than 90% of the trials at the key leading to the low-amount alternative (impulsiveness) whereas people chose the low-amount alternative on only 72% of the trials. This difference may reflect some fundamental difference between animal and human subjects, it may reflect a fundamental difference between probability and time or between the effective scales of these two dimensions, or it may have been due to the fact

that the low-amount reinforcer was almost *immediately* available to pigeons, but not *certainly* available to our human subjects ($p = 17/18$). Occasionally, subjects did not win the \$1 when they chose this alternative.

The matching law (Herrnstein, 1961) easily explains the shift in an organism's choices (B_1/B_2) when, as in the Rachlin and Green (1972) experiment, a constant delay (t) is added to two alternatives differing in amount (a_1 and a_2) and delay (d_1 and d_2):

$$\frac{B_1}{B_2} = \frac{a_1}{a_2} \cdot \frac{(t + d_2)}{(t + d_1)} \quad (1)$$

If the ratio of the amounts is less than the ratio of the delays, and t is small, the Equation 1 predicts that animals will prefer the alternative with the shorter delay. As t grows, however, the addition of a constant to both d_1 and d_2 brings the delay term of the matching equation closer to unity although the amount term is unchanged. Thus, regardless of how small the difference of amount originally was, at some point, as t increases, that difference will come to predominate, no matter how large the difference of delay originally was. As t increases, an animal that originally chose the alternative with the smaller delay ("impulsiveness") will come to choose the alternative with the larger amount ("self-control"). The matching law thus predicts that, with the commitment paradigm of Figure 1a, as t increases, an animal at B will come to choose C, commitment.

Although the matching law easily accounts for the results of the Rachlin-Green experiment in which time was varied, when probabilities are substituted for delays the matching law no longer accounts for the change in preference:

$$\frac{B_1}{B_2} = \frac{a_1}{a_2} \cdot \frac{(q \times p_1)}{(q \times p_2)} \quad (2)$$

The difference between Equation 1 and Equation 2 is that whereas delays must be added to calculate overall delay, probabilities must be multiplied to calculate overall (conditional) probability. Thus, in Equation 2, the value of q is irrelevant—the q 's cancel out—and the matching law cannot account for the shift in choice with different values of q . (This is just a restatement of the "independence axiom" of decision theory. The fact that q does affect

Fig. 3. Percentage of choices of larger reward at A and of commitment to larger reward at B for $q = 15/18$ and $q = 3/18$ (see Figure 1). The brackets indicate one standard deviation.

choice is one of the several violations of that axiom found in studies of choice; see Machina, 1987.) Because the concepts of matching and utility maximization both necessitate conformity to the matching law (Rachlin, Green, Kagel, & Battalio, 1976) and reflect our most fundamental conceptions of choice (Rachlin, 1971), the fact that Equation 2 does not explain the results of the current experiment is a matter of concern.

However, if we assume that delay, not probability, is the effective independent variable in humans' rule-governed choice as well as in animals' contingency-governed choice, then the present results are easily explainable by matching. If delay were the effective independent variable, then, where probabilities are concatenated, it would be proper to convert probabilities to delays and add the delays rather than multiply the probabilities.

In a series of trials with constant probability outcomes, the relationship between the arithmetic-mean-delay-to-outcome and probability-of-outcome is (from Rachlin et al., 1986):

$$d = \frac{i + c}{p} - i \quad (3)$$

where i is the interval between trials, c is trial duration, and p is the probability of the outcome. When $p = 1.0$ (certainty) the delay is

equal to a small constant (c). Thus, delay cannot actually equal zero. Otherwise the matching equation would predict infinite preference for certain rewards of vanishingly small but finite magnitudes. This corresponds to Mazur's (1987) device of adding a constant to actual delays in order to calculate preference via the matching equation.

Let us now attempt to roughly account for the present results using Equations 1 and 3. The average trial duration, the time between choice and outcome (c), was about 2 s, and the average intertrial interval, the time between the end of one trial and the start of the next trial (i), was about 13 s ($15 \text{ s} \times 20 \text{ trials} = 300 \text{ s} = 5 \text{ min} = \text{average session duration}$). For the several probabilities of this experiment, $3/18$, $4/18$, $5/18$, $15/18$, and $17/18$, the corresponding delays (by Equation 3) are 77 s, 55 s, 41 s, 5 s, and 3 s. For the choice at A:

$$\frac{B(\$1)}{B(\$4)} = \frac{1}{4} \cdot \frac{41}{3} = 3.4,$$

a strong preference for the small high-probability reward. For the choice at B, between the \$1 reward (usually chosen at A) and commitment to the \$5 reward ($q = 15/18$):

$$\frac{B(\$1)}{B(\$5)} = \frac{1}{5} \cdot \frac{(5 + 55)}{(5 + 3)} = 1.5,$$

still a preference for the small high-probability reward. However, for $q = 3/18$:

$$\frac{B(\$1)}{B(\$5)} = \frac{1}{5} \cdot \frac{(77 + 55)}{(77 + 3)} = 0.33,$$

a strong preference for the large low-probability reward. These calculations are only rough approximations for several reasons. First, Equation 3 gives the arithmetic mean delay of a probabilistic reward. The effective delay may be a geometric mean (Killeen, 1968) or some other measure. Second, the constants i and c used in Equation 3 are themselves approximations. Third, the matching equation was used in its unexponentiated simple form rather than its general form (Baum, 1974). Fourth, any effect of overall rate of reinforcement (in addition to that of delay) was ignored. Nevertheless, the conversion of probabilities to delays does explain the effect of q on choice.

Modern decision theory is concerned with how verbally presented probabilities are in-

ternally represented. The present behavioral analysis inverts this concern and asks: What do verbally (or pictorially) presented probabilities themselves represent? Once it is recognized that statements about probabilities represent delays (are discriminative stimuli for delay contingencies) some facts in the literature on human decision may be explained (cf. Rachlin *et al.*, 1986). For instance, the "certainty effect" (Kahneman & Tversky, 1984) is an overweighting of certainties (or very high probabilities) relative to uncertainties. Subjects' preference at A for the smaller more certain reward is an example of the certainty effect. The conversion of probabilities to delays in the matching equation explains the certainty effect as due to a nonlinear (hyperbolic) delay-discount function. Certainties represent immediacies (we claim) and immediacies are strongly overweighted (how much depends on how small the constant c is) by the hyperbolic function. Another unexplained effect of the decision literature is the "preference-reversal effect." When asked to *choose* between two rewards (called "prospects") differing in amount and probability, people tend to overweight probability (explained in our terms by the hyperbolic delay-discount function). But when they are asked to *evaluate* the prospects individually, people tend to overweight amount. For instance, with the prospects, $a = \$1$, $p = .5$ and $a = \$5$, $p = .1$, the former would be chosen over the latter in a simultaneous choice experiment but the latter would be evaluated higher individually. If probabilities actually represent delays, however, it may be that *evaluations* of probabilities represent additional delays. The response called for in an evaluation experiment is not a choice but a *judgment*. In behavioral terms, an expressed judgment is a discriminative stimulus for a choice. Hence, a judgment must precede a choice in time. (In cognitive theories a judgment is an output of an earlier stage in the decision process than is a choice.) A subject who evaluates an alternative produces a discriminative stimulus for some further action. Thus, evaluation adds a constant delay to the delays of the prospects being evaluated and, just as reversals occur with the commitment paradigm when t is increased (or q is decreased), so may reversals occur when people are asked to evaluate prospects rather than to actually choose among them.

CONCLUSIONS

A previous experiment reported in Rachlin et al. (1986), with human subjects and a spinner apparatus similar to the one used here, demonstrated a strong effect on probabilistic choice of intertrial-interval differences of less than 90 s. Neither that experiment nor the present one constitutes evidence for the existence of an internal mechanism that converts probabilities to delays. Both experiments do show that probabilities as stated or pictured in experiments with human subjects act like delays as programmed in experiments with animal subjects. Given this correspondence, the same terms should be used to describe both sorts of experiment. Rachlin et al. argued further that the language of delay is better than that of probability for such description because the language of delay fits more precisely with events as they take place over time in both the real world and in the laboratory. The present experiment adds to the case for describing events as delays rather than probabilities because delays combine additively (whereas probabilities combine multiplicatively) and the environmental variables in both human-probability and animal-delay choice experiments appear to combine additively as they act on behavior.

Does the relation between (a) humans' choices among probabilistic rewards and (b) animals' choices among delayed rewards extend to (c) humans' choices among delayed rewards and (d) animals' choices among probabilistic rewards? Although the Rachlin et al. (1986) experiment did find an interaction between probability and delay in human choice, it remains to be shown that our descriptive model fits categories (c) and (d). A problem in describing humans' choices among delayed rewards may be the effectiveness of the language of delay as a commitment device in itself (see Ainslie, 1975). A problem in describing animals' choices among probabilistic rewards may be the difficulties in specifying what constitutes a trial; this is one reason why Skinner abandoned mazes and alleys to investigate animal behavior in real time (Skinner, 1979). We hope, nevertheless, that the present analysis will provide a framework for eventual description of all four types of choice.

REFERENCES

- Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, **82**, 463-496.
- Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, **22**, 231-242.
- Hayes, S. C., Brownstein, A. J., Haas, J. R., & Greenway, D. E. (1986). Instructions, multiple schedules, and extinction: Distinguishing rule-governed from schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior*, **46**, 137-147.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, **4**, 267-272.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, **39**, 341-350.
- Killeen, P. (1968). On the measurement of reinforcement frequency in the study of preference. *Journal of the Experimental Analysis of Behavior*, **11**, 263-269.
- Machina, M. J. (1987). Decision-making in the presence of risk. *Science*, **236**, 537-543.
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 55-73). Hillsdale, NJ: Erlbaum.
- Mischel, W., & Grusec, J. (1967). Waiting for rewards and punishments: Effects of time and probability on choice. *Journal of Personality and Social Psychology*, **5**, 24-31.
- Navarick, D. J. (1987). Reinforcement probability and delay as determinants of human impulsiveness. *Psychological Record*, **37**, 219-225.
- Rachlin, H. (1971). On the tautology of the matching law. *Journal of the Experimental Analysis of Behavior*, **15**, 249-251.
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, **17**, 15-22.
- Rachlin, H., Green, L., Kagel, J. H., & Battalio, R. C. (1976). Economic demand theory and psychological studies of choice. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 129-154). New York: Academic Press.
- Rachlin, H., Logue, A. W., Gibbon, J., & Frankel, M. (1986). Cognition and behavior in studies of choice. *Psychological Review*, **93**, 33-45.
- Rotter, J. B. (1954). *Social learning and clinical psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Skinner, B. F. (1979). *The shaping of a behaviorist*. New York: Knopf.
- Stevenson, M. K. (1986). A discounting model for decisions with delayed positive or negative outcomes. *Journal of Experimental Psychology: General*, **115**, 131-154.

Received February 9, 1987
Final acceptance July 9, 1987