

*IMPULSE CONTROL IN PIGEONS*¹

G. W. AINSLIE

HARVARD UNIVERSITY

Pigeons were given a small, immediate food reinforcement for pecking a key, and a larger, delayed reinforcement for not pecking this key. Most subjects pecked the key on more than 95% of trials. However, when pecking a differently colored key at an earlier time prevented this option from becoming available, three of 10 subjects consistently pecked it, thereby forcing themselves to wait for the larger reward. They did not peck the earlier key when it did not prevent this option. This is an experimental example of psychological impulse and a learnable device to control it. Although only a minority of the subjects learned it, the fact that such learning is possible at all argues for a theory of delayed reward that can predict change of preference as a function of elapsing time.

The phenomenon of impulsiveness or failure to delay gratification is often discussed in sociology (Straus, 1962), economics (Strotz, 1956), and clinical psychology (Freud, 1958; Mischel and Metzner, 1962). An impulse can be defined as the choice of a small, short-term gain at the expense of a large, long-term loss. It often is described by such terms as spendthriftiness, temptation, or seduction. Although many impulsive behaviors involve the taking of risks or the weighing of a reward against a punishment, their common characteristic is that they trade a high overall expectancy of gain for a low one. Mowrer and Ullman (1945) advanced the hypothesis that impulsiveness arises from a basic property of reward that can be observed in lower animals. This property is a decline in the effectiveness of a reinforcement as it is delayed from the moment of choice, and is usually described as a steep, negatively accelerated curve (Kimble, 1961, pp. 140-144; Renner, 1964). Such a curve suggests that small, early reinforcements may have a great advantage over much larger, delayed ones, and raises the question of how impulsive behavior is ever prevented.

The answer may lie in the shape of the delay function. In the last decade, several para-

metric experiments dealing with preference for various amounts of reinforcement at various delays have generated highly concave curves of reinforcing effect as a function of delay (Chung and Herrnstein, 1967; Killeen, 1970; Logan, 1965; Shimp, 1969). One prediction that can be made from these curves is that preference for some alternative reinforcements at different delays can change simply as a function of elapsing time. For instance, Chung and Herrnstein (1967) found that food reinforcements were preferred in inverse proportion to their delay. It would follow that if a reinforcement were available at time T , and an alternative reinforcement that was three times as great were available at $T + 3$ sec, a subject making the choice at $T - 3$ sec would choose the larger, later reinforcement, while if it chose at $T - 1$ sec it would choose the smaller, earlier one.

If there exist cases where preference changes simply as a function of the time the choice is made, they pose for a subject the problem of binding its own future freedom of choice. Any effective device for getting the later, larger reinforcement must include a means of either preventing preference from changing as the smaller, earlier reward comes close, or keeping the subject from acting on this change. If the subject lacked the ability to do this on its own, it could be expected to learn a device to do it if an experimenter offered one. This prediction would not depend on the existence of "higher intelligence", but only on whether the effect of the larger reinforcement was great enough

¹The work reported here has been supported by a grant to Harvard University from the National Institute of Mental Health (MH-15494-04). I thank Richard Herrnstein and Howard Rachlin for advice and Donald Loveland for technical assistance. Reprints may be obtained from G. W. Ainslie, Massachusetts Mental Health Center, 74 Fenwood Rd., Boston, Mass. 02115.

to cause learning of the pre-committing device at the time the device was available. If the subject's preference did not change regularly over time, a device to bind the behavior would not be differentially reinforced and thus would never be learned.

The author has previously found evidence that pigeons can learn a pre-committing device (described in Rachlin, 1970). Subjects could avoid the option of getting a reduced, immediate food reinforcement by pecking a key several seconds before the option was available. But only two birds did this out of 10 tested. Rachlin and Green (1972) were able to increase this proportion by presenting pigeons with two keys, a committing and a non-committing key, the consequences being determined by whichever key received the bird's twenty-fifth peck. However, interpretation of this experiment is complicated by the fact that the time required by most of the birds to peck 25 times was much longer than the longest possible delay to which they might commit themselves, so that to obtain "immediate" reinforcement they would have to start pecking a long time before this reward was actually available. In order to study the pre-committing response in its most elementary form, the present experiment again used discrete trials with a single key. Because a large number of birds could not be run for the necessary length of time, potential subjects were screened to find those that learned pre-commitment, and those that did were studied extensively.

METHOD

Subjects

Ten White Carneaux pigeons, one of which had been in the most recent pilot study and nine previously naive birds that had just been trained to peck a red key for food with every peck reinforced, were kept at 80% of their free-feeding weights.

Apparatus

The subjects performed in a sound-proof chamber that measured 30 by 32 by 33 cm. In one wall was a single key, which was centered at a height of 24 cm and required 14 g (1.4 N) to operate. The key could be illuminated from behind by green or red 7-W Christmas-tree bulbs. Ten cm beneath the key was a 5 by 6

cm niche in the wall, into which a hopper of grain could be raised by an electromagnet. Centered in the top of the chamber was a one-way viewing lens, of the kind often set into the front door of apartments. It measured 1 cm in diameter and permitted the experimenter to see most of the chamber without being seen from inside. No houselight was used. White noise was piped into the box to mask environmental sounds.

General Procedure

Each trial lasted 19 sec, regardless of the subject's behavior. Each subject had 50 trials a day. Each subject served as its own control. Seven subjects that did not meet criteria described below were eliminated. The three remaining subjects, including the bird from the pilot study, each ran in three successive experiments that differed from one another only in the kind of control condition used (Figure 1). The birds went twice from the experimental condition to each control condition and back again, for a total of 12 changes over a period of 23 months. The temporal parameters used were those that had been the most productive in previous pilot experiments.

Experimental condition. At the beginning of each trial, the key was illuminated green

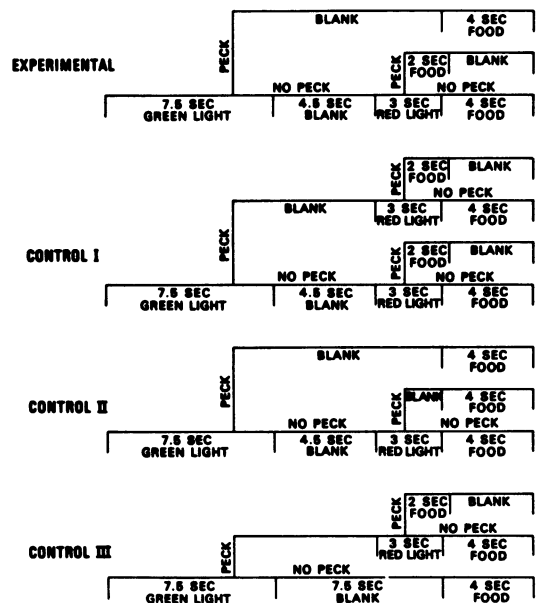


Fig. 1. Diagram of the conditions. Events proceed in a horizontal line. A response caused a vertical move to a different horizontal line.

for 7.5 sec. If it was not pecked, the key went dark after the 7.5 sec and remained so for another 4.5 sec, after which it was illuminated red for 3 sec. If it was still not pecked, it went dark at the end of the 3 sec and the subject was given access to grain for the next 4 sec, which were the last 4 sec of the trial. Pecking the key while it was green made the key go dark and prevented it from being lit red later in the trial. The subject was then given access to grain for the last 4 sec of the trial. Pecking the key while it was red made it go dark and gave the subject access to grain for the next 2 sec. The key then remained dark and the subject had no further access to grain for the rest of the trial. Since it takes pigeons about 1 sec to get to the grain in this kind of apparatus, the actual reinforcements were probably about 3 sec *versus* 1 sec of feeding.

Control I. This was the same as the experimental condition, except that a peck on the key while it was green did not prevent it from being lit red later in the trial. Thus, it had no pre-committing effect. A subject was shifted to the Control I when it had pecked the green key on more than 40% of trials for three consecutive five-day periods. It went from the Control I to the experimental condition when it had pecked the green key on fewer than 30% of trials for three consecutive five-day periods. Two birds that pecked the red key on fewer than 90% of trials, four birds that never pecked the green key enough to be shifted into Control I, and one bird that never pecked the green key on fewer than 95% of trials were discarded after 3000 trials (Figure 2, bottom row).

Control II. This was the same as the experimental condition, except that a peck on the key while it was red had no effect, other than to darken the key. Subjects thus always received the later, larger reinforcement. This control served the same purpose as Control I, except that it would also rule out the possibility that the birds pecked the green key just to control the appearance of the red key *per se*. Beginning with the first run of Control II, all birds were given 40 trials a day for 48 days before the condition was changed, regardless of their behavior. The reduction in the number of trials per day was made to prevent the birds from gaining too much weight, since they had 4-sec access to food on every trial. The number of days was fixed to eliminate any

question of artifact arising from long-term, random changes in their green-key pecking.

Control III. This was the same as the experimental condition, except that a peck on the key while it was green was *necessary* for the key to be lit red later in that trial. Thus, not pecking the green key in Control III had the same effect as pecking it in the experimental condition. This condition was designed to find whether choice was changing haphazardly during the period the key was green, causing pecking whenever that alternative was transiently preferred.

RESULTS AND DISCUSSION

Eight of the 10 original birds initially pecked the key when it was red on more than 95% of trials. They all continued to peck it on virtually all the trials in which it appeared for as long as they remained in the experiment. During Control II, when the key lit-up red on all trials, no bird pecked it on fewer than 95% of trials in any six-day period. During the experimental condition and in Control I and Control III, when the appearance of the red key depended on the subject's previous behavior, no subject pecked it on fewer than 95% of the trials in which it appeared in any six-day period (or five-day period, before the number of trials in a day had been reduced to 40). Thus, the tendency to seek the shorter access to food when it was immediately available remained strong after as many as 20,000 trials. However, casual observation through the one-way viewing lens often revealed the subjects to be pecking the chamber wall near the key while it was red, before pecking the key itself. This may be an example of the "mediating behavior" that has been described during differential-reinforcement-of-low-rate responding (DRL: Bower, 1961; Kramer and Rilling, 1970).

The percentage of trials on which subjects pecked the key when it was green is shown in Figure 2. Each of the three subjects that initially pecked it on a large proportion of trials during the experimental condition and on a small proportion of trials during Control I maintained or increased this difference between experimental and control conditions throughout the experiment. During the experimental condition, when pecking the green key was required to avoid the option of getting

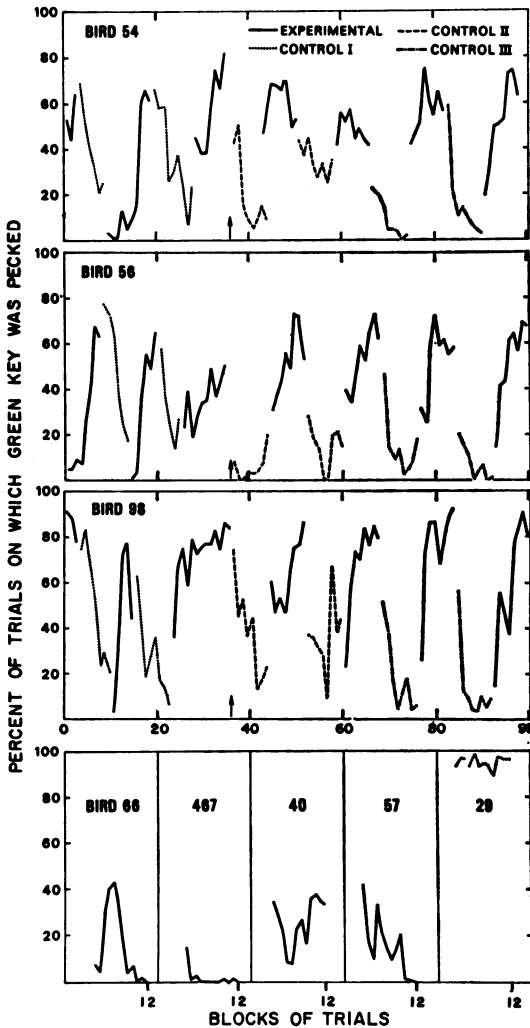


Fig. 2. Percentage of trials on which subjects pecked the key while it was green, for each five-day period (250 trials; six-day periods, containing 240 trials, after arrow). Data from the two subjects that did not reach criterion response rates on the red key are omitted.

the shorter, earlier access to food, this pecking rose to between 50 and 90% of trials. During Control I and Control II, when pecking the green key did not affect the occurrence of this option, pecking the green key fell to much lower levels. A similar fall was seen during Control III, when pecking the green key was required in order for the option of getting the shorter, earlier access to food to occur.

Thus, three subjects worked at different points in the trial to obtain and to avoid the same event. However, the food itself may not have been the only important reinforcing ele-

ment of this event. Recent experiments have made it clear that pigeons have a strong tendency to peck a key that once has been associated with food, even when it currently reduces or eliminates their food intake (Fantino, 1966; Gamzu and Williams, 1973; Williams and Williams, 1969). Therefore, it is not possible to say what part of the reinforcement for pecking the key when it was red came from the food itself, and how much from the chance to peck a food-related key. The attractiveness of the earlier reinforcement may have been augmented by the chance to peck a key to get it, or changed in either direction by the recent consumption of food during the preceding trial. What is remarkable is that some pigeons learned to peck a key at an earlier time, if and only if this made them unable to obtain the smaller reinforcement. At one point in the cycle, these subjects had a great tendency to peck the key for an early reinforcement, but at a greater distance they worked to forestall this tendency. This seems to be a true example of impulsiveness and impulse control as the terms are generally used.

Preference for pecking the key may not have been constant during the 3-sec period it was red. The observation of mediating behaviors during this time suggests that it was not, since these presumably were attempts to withhold key-pecking behavior. But since any change of preference while the key is red confronts the subject with the problem of impulse control, this distinction is not important. In either case, if the subject prefers the later, larger reinforcement at first, it will need to adopt some kind of pre-commitment in order to get it.

The majority of subjects might have failed to learn the pre-committing device because the effectiveness of the reinforcements used declined in such a way that preference never changed. However, it may be that the differential effectiveness of the reinforcements during the time the key was green was too small to allow most pigeons to learn a pecking task; or these contingencies of reward may have been too confusing for them. No conclusions can be drawn from the negative findings.

This experiment affirms and enlarges on the view of impulse control advanced theoretically by the economist Strotz (1956), and a related proposal by the sociologist Becker (1960). The existence of an experimental paradigm of

impulse and control may permit research on a subject that has been almost purely theoretical in the past.

REFERENCES

- Becker, H. S. Notes on the concept of commitment. *American Journal of Sociology*, 1960, **66**, 32-40.
- Bower, G. H. Correlated delay of reinforcement. *Journal of Comparative and Physiological Psychology*, 1961, **54**, 196-203.
- Chung, S. H. and Herrnstein, R. J. Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1967, **10**, 67-74.
- Fantino, E. Immediate reward followed by extinction vs. later reward without extinction. *Psychonomic Science*, 1966, **6**, 233-234.
- Freud, S. Formulations on the two principles of mental functioning. In J. Strachey and A. Freud (Eds.), *The standard edition of the complete psychological works of Sigmund Freud*. London: Hogarth, 1958. V. 12, Pp. 218-226.
- Gamzu, E. R. and Williams, D. R. Associative factors underlying the pigeon's key pecking in auto-shaping procedures. *Journal of the Experimental Analysis of Behavior*, 1973, **19**, 225-232.
- Killeen, P. Preference for fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1970, **14**, 127-131.
- Kimble, G. *Hilgard and Marquis' conditioning and learning*. New York: Appleton-Century-Crofts, 1961. Pp. 140-144.
- Kramer, T. J. and Rilling, M. Differential reinforcement of low rates: a selective critique. *Psychological Bulletin*, 1970, **74**, 225-254.
- Logan, F. A. Decision making by rats: delay versus amount of reward. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 1-12.
- Mischel, W. and Metzner, R. Preference for delayed reward as a function of age, intelligence, and length of delay interval. *Journal of Abnormal and Social Psychology*, 1962, **64**, 425-431.
- Mowrer, O. H. and Ullman, A. D. Time as a determinant in integrative learning. *Psychological Review*, 1945, **52**, 61-90.
- Rachlin, H. *Introduction to modern behaviorism*. San Francisco: Freeman, 1970. Pp. 186-191.
- Rachlin, H. and Green, L. Commitment, choice, and self-control. *Journal of the Experimental Analysis of Behavior*, 1972, **17**, 15-22.
- Renner, K. E. Delay of reinforcement: a historical review. *Psychological Bulletin*, 1964, **61**, 341-361.
- Shimp, C. P. The concurrent reinforcement of two interresponse times: the relative frequency of an interresponse time equals its relative harmonic length. *Journal of the Experimental Analysis of Behavior*, 1969, **12**, 403-411.
- Straus, M. A. Deferred gratification, social class, and the achievement syndrome. *American Sociological Review*, 1962, **27**, 326-335.
- Strotz, R. H. Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 1956, **23**, 165-180.
- Williams, D. R. and Williams, H. Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior*, 1969, **12**, 511-520.

Received 4 September 1973.

(Final Acceptance 10 December 1973.)