Behav. Res. & Therapy. 1974. Vol. 12, pp. 89–97. Pergamon Press. Printed in Great Britain

MAINTENANCE AND TRANSFER OF SELF-REINFORCEMENT FUNCTIONS*

ALBERT BANDURA and MICHAEL J. MAHONEY

Department of Psychology Stanford University, Stanford, California, U.S.A.

(*Received* 11 *September* 1973)

Summary—In order to examine conditions maintaining self-reinforcement functions, pigeons were trained to reward their own performances and then tested for adherence to work requirements under decreasing likelihood of punishment for undeserved self-reward. Contingent self-reinforcement was stably maintained given moderate to high probability that unmerited self-reward would incur punishing consequences. In successive reversals of treatment and test conditions, the amount of behavior performed for each self-reward covaried with self-reinforcement rate. A further experiment demonstrated that self-reinforcing practices, involving both performance and consummatory contingencies, transfer to new activities for which the animal had never been trained to reward himself contingently. Adherence to performance requirements was more stringent, however, than to limitations on amount of reinforcers consumed from freely available provisions.

It was previously shown that self-reinforcing functions, which play an influential role in the regulation of human behavior, can be established in animals as well (Mahoney and Bandura, 1972). In this process, animals maintain effortful behavior by administering to themselves reinforcers which they control whenever they attain criterion performances. This research further reveals that animals will adhere to performance requirements for self-reward for some time after unmerited self-reinforcement is no longer punished. However, after experiencing that they can safely self-feed without prior work, animals eventually discard performance contingencies. The more onerous the work requirements, the less enduring are the self-imposed contingencies.

In the social learning conditions of everyday life, it is rare for undeserved self-reward to go completely ignored. People who reward themselves for inadequate performances are likely to be rebuked at least on some occasions. Intermittent punishment for unmerited self-reward has, in fact, been proposed as one of several conditions supporting self-reinforcement systems (Bandura, 1971). The first experiment reported in this article investigated the role of punishment in sustaining contingent self-reinforcing practices.

Pigeons were trained to self-impose a high performance output for each self-reward and then were tested for adherence to work requirements when unwarranted self-feeding was no longer punished. After several reversals of these procedures, maintenance of self-imposed contingencies was measured under progressively diminishing rates of punishment for transgressive self-feeding. Variable intermittent punishment of undeserving self-reward was expected to increase the endurance of self-reinforcement systems.

* This research was supported by Research Grant M–5162 from the National Institute of Mental Health.

METHOD

Subjects

Two male White King pigeons, who were experimentally naive, served as subjects. Both birds were deprived to 75–80 per cent of their free-feeding weight throughout the study.

Apparatus

A standard Lehigh Valley pigeon chamber was equipped with a photocell for monitoring feeding behavior (Mahoney and Bandura, 1972). Grain reinforcement was presented via a food hopper. A response disc located directly above the hopper entrance was illuminated red until a fixed number of key pecks had been performed, whereupon the illumination changed to white signifying attainment of performance requirements for self-reward. The response disc was neither illuminated nor operative during intervals between trials.

Training and test procedures

Self-reinforcement patterns were established using the training procedures described by Mahoney and Bandura (1972). Initially, the birds had to peck the response key in order to produce the hopper filled with mixed grain. Gradually, the hopper was presented at earlier stages of the pecking sequence until eventually it appeared before the bird performed any pecking responses. If the bird entered the hopper without performing the required pecking response the chamber was darkened and the food was immediately removed for 10 sec. In this way, the birds learned to peck before eating.

Daily training sessions were conducted until the birds completed 50 reinforced trials. Each trial was initiated by presentation of the hopper filled with grain. If the bird pecked the disc before entering the hopper, he was allowed 2 sec access to the grain. After the feeding period was completed, the hopper was withdrawn and all lights were turned off for 10 sec, which was the programmed interval between trials. If, on the other hand, the bird entered the hopper without first pecking the disc, the hopper was immediately withdrawn and the lights were turned off for the duration of the intertrial interval.

The birds were first trained to feed themselves for performing one pecking response (FR1). After this self-reinforcement rate stabilized at the 80 per cent level, the performance criterion was successively raised by one response whenever they attained an 80 per cent self-reinforcement rate on the preceding ratio schedule.

When the birds reached the 80 per cent self-reward rate for the FR5 schedule, punishment of unmerited self-feeding was discontinued to test for maintenance of the self-reinforcement system. During the test trials, the achievement light cue appeared when the appropriate number of responses was completed, but the birds could self-feed without pecking and receive no negative consequences.

After two days of testing on the FR5 schedule the performance standard was progressively raised by one response until the birds discarded the work contingency for self-reward. When this occurred the training conditions were reinstated until contingent self-reinforcement was re-established. The standard was then successively raised to 150 per cent of the test performance level at which the animals had abandoned the work requirement for self-reward. The higher performance demands were programmed in order to permit a stringent test of maintenance of onerous outputs that exceeded the animals' prior tolerances. When they displayed a self-reinforcement rate above 80 per cent on two consecutive days they were again tested without punishment for unmerited self-reward until they discarded the

performance requirement. After the second test, the animals were again trained to reward themselves for attaining performances at the highest previous training level.

Maintenance conditions

During the initial maintenance conditions, each transgression (i.e. self-feeding without working) was punished by food withdrawal. When the birds had maintained a similar self-reinforcement rate for two consecutive days, the probability of punishment for transgressions was decreased by 10 per cent. The likelihood of punishment was thus progressively reduced until eventually it was eliminated completely.

RESULTS

The self-reinforcement rates during training, testing, and maintenance phases of the experiment are summarized graphically in Fig. 1 and 2. Both animals adopted increasingly higher performance standards during training but discarded them after negative sanctions for undeserving self-reward were removed. Interestingly, during initial test sessions, transgressive self-feeding on early trials in the session typically resulted in stringent adherence to performance requirements on later trials. Such response changes, which do not follow an extinction course, suggest that transgressive behavior may have acquired through prior punishment sufficient threat value to reinstate temporarily contingent self-reward practices.

The influential role of intermittent punishment in sustaining self-reinforcement systems is revealed in the results from the maintenance condition. Animals adhered to stringent performance requirements before rewarding themselves for nearly a thousand trials as long as there was a 50 per cent chance or better of being punished for unmerited self-reward. However, less probable punishments had relatively weak sustaining value.

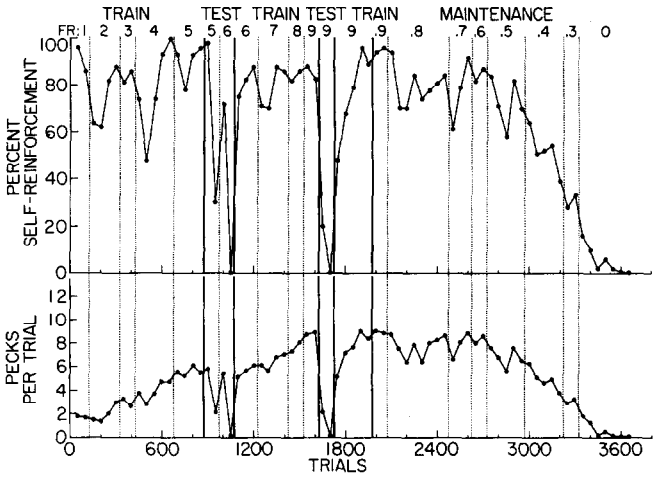

Fig. 1. Rate of self-reinforcement and number of responses performed per self-reward during training, test, and maintenance conditions. The whole numbers at the top of the graph designate the fixed-ratio schedules in effect during a given period. All data points in the test trials represent contingent self-reward in the absence of punishment for transgressions. The decimalized numbers in the maintenance phase indicate the proportion of noncontingent self-rewarding responses that were punished.

Results on the number of responses the birds performed before engaging in self-feeding closely parallel the rate of self-reinforcement. During periods when the animals reinforced themselves contingently they also performed a large number of responses for each self-reward. As the rate of contingent self-reinforcement declined, on the trials in which the animals adhered to performance requirements, they executed decreasing amounts of behavior before feeding themselves.

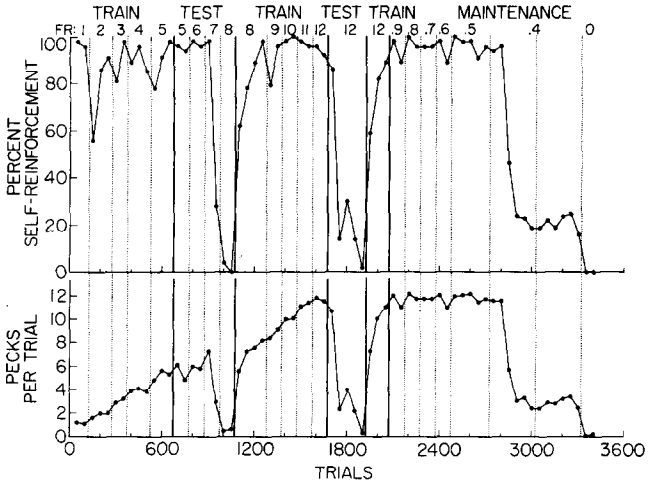

Fig. 2. Rate of self-reinforcement and number of responses performed per self-reward during training, test, and maintenance conditions.

Self-maintenance of anomalous performances

Because of the inherent relationship between pecking and grain reinforcement, a second experiment was conducted to test the generality of self-reinforcement functions when the required performances are arbitrary and unnatural. A male White King pigeon was administered the identical procedures used in the preceding experiment except that the requisite performance involved pressing a treadle rather than pecking a disc.

A treadle was installed on the floor of the chamber four inches to the right of the hopper opening. The response required the pigeon to raise a foot off the floor, exert moderate pressure on the treadle and then release it. This proved to be an awkward and difficult task for the bird because raising his foot threw him off balance. Although the panel response disc was inoperative, it remained illuminated red until a treadle press occurred, whereupon it turned white. After being trained to reward himself with food for performing at least one treadle response, the animal was tested for adherence to work requirements under nonpunishment conditions, retrained again, and then placed on maintenance conditions with decreasing likelihood of punishment for noncontingent self-feeding.

As shown in Fig. 3, the results obtained with pecking responses are essentially replicated with laborious treadle responses. Self-imposed performance requirements were readily adopted, they were later discarded in the absence of negative sanctions, but they were well maintained under moderate to high likelihood of punishment for unmerited self-reward.

Development and transfer of dual forms of self-regulation

In the experiments discussed thus far animals adopted performance controls on self-feeding but consumed whatever food was available. The present study sought to establish a more taxing form of self-regulation in which animals control the amount of food consumed from a large available supply as well as adhere to work requirements for self-feeding. This dual self-regulation corresponds more closely to human forms of self-control (Thoresen and Mahoney, 1974).

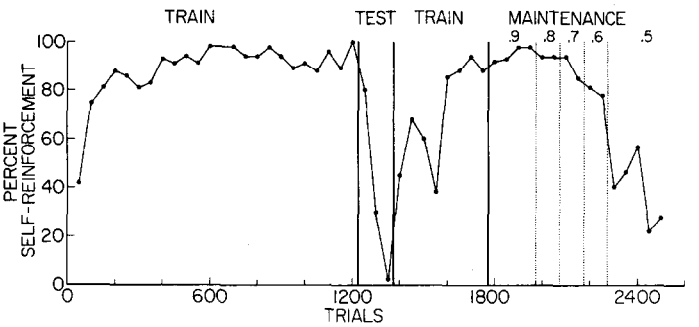

Fig. 3. Rate of self-reinforcement and number of responses performed per self-reward during training, test, and maintenance conditions.

Development of performance standards and self-reinforcing functions would have limited value if the effects of training never generalized beyond the specific task. It is assumed that, in the absence of differential standard setting, after the evaluative properties of given accomplishments are established, performances that exceed or fall below acceptable levels are likely to elicit similar self-reinforcing responses on different activities (Bandura, 1969). That is, a person for whom average performances have been repeatedly devalued will respond to modal achievements on new tasks as undeserving of self-reward, and attainments that surpass modal levels as commendable.

The present experiment was designed further to test whether an animal would impose both performance and consummatory contingencies for self-reward on new tasks that differed from the one on which the self-reinforcing system was originally established.

Subject

A four-year-old male Cocker-Poodle served as the subject. As a household pet he had been previously trained to perform several responses such as rolling over, sitting up, and barking on verbal cue. Prior to the formal experiment, the remark 'good dog' was established as a secondary reinforcer through association with food, and the cue 'bad dog' as a negative reinforcer through pairing with food removal and time-out from food reinforcement. These verbal sanctions were used to establish self-reinforcing functions. The dog was deprived of food for two hours prior to each session.

Apparatus

The experiment was conducted in a sound-proof room equipped with a one-way mirror and an intercom system. The manipulanda consisted of an IBM Selectric typewriter, a telegraph key, a plastic hoop, and a foot treadle.

Development of self-reinforcing functions

The dog was initially trained by successive approximations and food reinforcement to stand on his hind legs and to strike the keyboard of an electric typewriter located on a typing stand. Typing was defined as placing one or both paws on the typewriter keyboard and depressing at least one key with sufficient force to produce an audible noise and a typed character. Performances during all training and testing sessions were recorded by an independent scorer who observed the sessions through a one-way mirror in an adjoining room. Meat-flavored cubes served as the self-rewards for typing performances.

After establishing the typing response, a meat cube was placed on a white card on the floor 18 in. in front of the typewriter. If the dog approached or ate the meat cube prior to typing, the experimenter said 'bad dog' and imposed a 10 sec time-out by instructing the dog to lie down at the far end of the room. When the dog typed before feeding himself, the experimenter said 'good dog'. This training procedure was gradually extended by adding meat-laden cards so that the dog had to return to the typewriter after each self-feeding despite the continued presence of other quantities of food. Placement of meat cubes was initially spread out, then gradually brought closer together. The successive addition of food was continued until 10 meat-laden cards were set around the typewriter in close proximity to one another.

After the dog achieved a session of complete self-control by typing before self-feeding and treating himself to only one piece of food for each typing performance, the experimenter conducted the further training sessions from the observation room using the intercom system. On transgression trials the dog was instructed to lie down in the time-out area. If he failed to do so, which happened only twice, the experimenter returned to the room and repeated the instruction until it was followed.

It required approximately half an hour of shaping to get the animal to type and then to reward himself. Because of the speed with which self-reinforcing functions were acquired it was decided to continue the training sessions to ensure that the system of self-reinforcement was stable before instituting test conditions. Five sessions, consisting of 10 self-reinforcements per session, were conducted each day. After the self-reinforcing responses were well established, the animal was tested for adherence to the adopted contingencies in the absence of the experimeter or verbal reprimands for transgressive self-feeding. During each session the dog was placed alone in the room along with the typewriter and 10 separate piles of food. The observer recorded the incidence of transgressive self-feeding and amount eaten on each trial.

Transfer of self-reinforcing functions

The transfer experiment was conducted following complete extinction of the self-reinforcement pattern on the typing task. Training in contingent self-reward with achievement cues was conducted on a new activity. The dog was reinforced with food for pressing a telegraph key. A key press also produced a tone which served as the performance achievement cue for self-reward. External reinforcement was gradually reduced to a 10 per cent schedule as key press performances were brought under self-reinforcement control through the same procedures by which the animal learned to self-reward typing performances.

Training in self-reinforcement for key presses was alternated with sessions in which the animal was reinforced with food for performing the first transfer task, which consisted of

jumping through a 30-in. hoop raised four inches off the floor. After the jumping response had been mastered, two sessions of 10 trials each were conducted in which hoop jumping produced the achievement tone but no rewards of any kind. The purpose of these trials was to demonstrate that jumping performances generate the attainment cue as well. In order to prevent extinction of jumping responses during this period, the two sets of cue generating trials were separated by a session of 10 trials in which jumping without the accompanying tones was externally reinforced.

A second transfer test was conducted with treadle performances. The treadle responses were trained and performed in attainment cue generating sessions without reinforcement in the same manner as the jumping responses.

In testing for transfer of performance and consummatory contingencies for self-reward, the dog was placed in the room alone with the upright hoop and 10 separate pieces of food for three consecutive sessions. Identical test procedures were used with the treadle transfer task. The observer recorded whether the animal performed jumping and treadle responses before self-feeding and the amount consumed during each self-feeding.

RESULTS

The per cent of trials that the animal adhered to performance and consummatory contingencies on the typing task is summarized graphically in Fig. 4. During the first 250 test trials he regularly typed before each self-feeding and ate only one meat cube per trial even though the food was freely available without any restrictions or negative sanctions. In the subsequent 50 trials he began to discard the contingencies, and shortly thereafter, he ate generously without performing any typing behavior.

Figure 5 shows the degree to which self-reinforcing functions transferred to new activities. Although the animal had never been trained to reward himself contingently on the

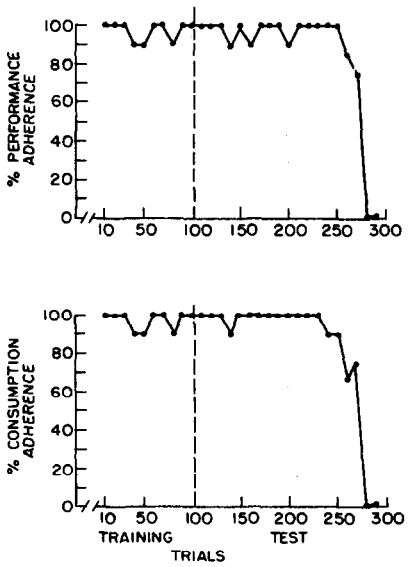

Fig. 4. The top figure shows the per cent of occasions on which the animal typed before eating; the lower figure indicates the percentage of times the animal ate only one piece of food from the freely available supply.

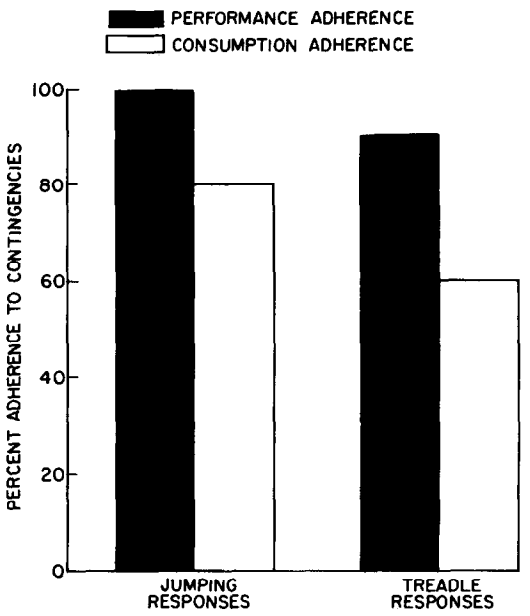

Fig. 5. The relative frequency with which the animal adhered to performance and consummatory contingencies on the transfer tasks.

transfer tasks, nevertheless, he self-imposed performance requirements without exception on the jumping task and he rarely ate without first performing treadle responses. He was somewhat less stringent, however, in adhering to a fixed amount of self-reward.

DISCUSSION

Findings of this series of experiments demonstrate a striking consistency across subjects, responses, and species in the acquisition and maintenance of self-reinforcement functions. Animals can be taught to reward their own performances, but they are quick to discard self-imposed work requirements if unmerited self-reward goes unpunished.

Results from the maintenance procedures support the proposition that intermittent punishment of undeserved self-reward plays a noteworthy role in maintaining self-reinforcement functions. As long as there existed a moderate to high likelihood of being punished for noncontingent self-reward, animals continued to adhere to performance requirements before reinforcing themselves. Given reasonable certainty of punishment, the negative sanctions rarely had to be used to ensure optimal self-maintenance of behavior through contingent self-reward. During the first thousand trials, when sanctions were periodically in effect, time-out punishment was actually administered on only two to three per cent of the occasions. Under low likelihood of punishment, however, high transgression rates produced more occasions on which punishment was administered but its potential influence was negated by the frequent experience of transgressive self-reward without negative consequences.

The combined results of several experiments suggest that the optimal level of sanctions for maintaining self-reinforcement functions depends upon the onerousness of the requisite performances. In an earlier study (Mahoney and Bandura, 1972), animals that adopted on

their own a relatively high performance output for each self-reward abandoned self-imposed contingencies more rapidly than did animals pursuing a minimal work schedule. In the present experiment, treadle responses, which were more arduous than pecking, required a higher rate of punishment to support self-reward on a performance-contingent basis. The suggested relationship between adherence to burdensome performance standards and amount of punishment supports would seem to warrant systematic examination.

Successful replication of self-regulatory phenomena with markedly different responses increases the generality of the findings. Data from the treadle task additionally refute the notion that birds might peck discs because such responses are inherently reinforcing. Although the illuminated disc was present, pecking responses, which had no functional value in this particular situation, were never performed. Given further evidence that birds cease pecking discs when self-feeding is permissible without prior performance of such responses, variations in pecking performances are better explained by changes in reinforcement conditions than by the properties of the behavior itself.

The experiment demonstrating transfer of self-reinforcement functions to new activities involved self-regulatory phenomena that bear some resemblance to human functioning. Performances were self-maintained over a continuing period without external intervention through regulated use of reinforcers from a freely available supply. Since food exercises powerful control over eating behavior, it is not surprising that adherence to performance requirements was more faithful than adherence to consummatory limitations. To interrupt eating for work when faced with appetizing provisions in the absence of social controls is a taxing order.

A rudimentary performance standard was adopted as the discriminative stimulus for self-reward. In the human case, self-reinforcement is usually tied to differential attainment levels as designated by scores, grades, or evident qualitative features signifying excellence. Performances that match or exceed adopted standards of merit serve as occasions for self-reward, whereas those that fall short of valued levels elicit negative self-reactions. There is every reason to expect that self-reinforcing practices in animals can be similarly linked to qualitative performance values through differential reinforcement.

REFERENCES

BANDURA A. (1969) *Principles of Behavior Modification.* Holt, Rinehart & Winston, New York.

BANDURA A. (1971) Vicarious and self-reinforcement processes. In *The Nature of Reinforcement* (Ed. R. GLASER), Academic Press, New York.

MAHONEY M. J. and BANDURA A. (1972) Self-reinforcement in pigeons. *Learning and Motivation* **3**, 293–303.

THORESEN C. E. and MAHONEY M. J. (1974) *Behavioral Self-control.* Holt, Rinehart & Winston, New York.