J. Behav. Ther. & Exp. Psychiat. Vol. 15, No. 4, pp. 291–298, 1984.
Printed in Great Britain.

0005-7916/84 $3.00 + 0.00

AN EFFECT OF IMMEDIATE REINFORCEMENT AND DELAYED PUNISHMENT, WITH POSSIBLE IMPLICATIONS FOR SELF-CONTROL*

ROBERT EPSTEIN

Harvard University

Summary—Behavior said to show self-control occurs virtually always as an alternative to behavior that produces conflicting consequences. One class of such consequences, immediate reinforcement and delayed punishment, is especially pervasive. Three experiments are described in which an effect of immediate reinforcement and delayed punishment is demonstrated. The results suggest that when immediate reinforcement and delayed punishment are imminent, the reinforcer alone controls the organism's behavior (in other words the organism behaves "impulsively"). The key to self-control, therefore, may be the acquisition of a large number of avoidance behaviors relevant to reinforcers that are correlated with delayed punishment. Human self-control may indeed involve such a process but undoubtedly involves others as well.

Human behavior often produces more than one consequence. The case in which consequences are conflicting—that is, in which behavior produces both reinforcing and punishing consequences—has been of special interest to both therapists and researchers for many years. On the one hand, conflicting consequences have been associated with various experimentally-induced behavioral abnormalities, such as "experimental neurosis" (Masserman, 1943; cf. Gantt, 1944) and approach-avoidance behavior (Miller, 1937, 1959). On the other hand, they have formed an integral part of research on punishment, since the behavior being punished is typically maintained by contemporaneous reinforcement [see Azrin and Holz (1966) for examples].

In either case, researchers have dealt primarily with behavior producing simultaneous reinforcement and punishment (e.g. Masserman, 1943) or with behavior producing reinforcement on some occasions and punishment on others (e.g. Miller, 1937; cf. Holz, 1968). Relatively neglected has been a class of conflicting consequences which is both pervasive and troublesome for human behavior, namely immediate reinforcement followed by delayed punishment. Cigarette smoking, overeating and drug abuse all produce such consequences and, perhaps for that reason, have been resistant to treatment (e.g. Stunkard, 1977). Many less insidious behaviors, such as classroom misbehaviors, also produce such consequences.

The present paper describes an effect of immediate reinforcement and delayed punishment on the behavior of pigeons and suggests an interpretation in terms of current conceptions of human self-control.

EXPERIMENT 1

METHOD

Subjects

Three experimentally naive, adult male White Carneaux pigeons, numbers 398WP, 233WP and 399WP, served as subjects. The birds were maintained at approx. 80% of

*Based on a chapter of a Ph.D. thesis presented to Harvard University. Portions were presented at the 87th annual meeting of the American Psychological Association, New York, New York, September 1979, the 51st annual meeting of the Eastern Psychological Association, Hartford, April 1980, and the 5th Harvard Symposium on Quantitative Analyses of Behavior, Cambridge, June 1982. The research was supported by Grants MH-15494 from the National Institute of Mental Health and BNS-8007342 from the National Science Foundation.

Requests for reprints should be addressed to Robert Epstein, Cambridge Center for Behavioral Studies, 11 Ware Street, Cambridge, MA 02138, U.S.A.

their free-feeding weights. Before the start of the experimental conditions each bird was wired for shock delivery according to the method described by Azrin (1959).

Apparatus

A standard two-key pigeon chamber was used. The two keys were on one wall at a height of 23 cm and spaced 13 cm apart. The left key could be transilluminated from behind with yellow light and the right, with red light. The pigeons could be given access to mixed grain through a recess in the wall beneath the keys. The chamber was equipped with overhead white house lights. A mercury commutator was attached over a 2-in. hole in the center of the ceiling of the chamber, and from the commutator was suspended a wire that could be plugged into a harness on each pigeon's back and through which electric shocks could be delivered. The chamber was enclosed in a sound-attenuating box, which was equipped with a ventilating fan as well as a speaker that emitted white noise.

Procedure

Reliable key pecking was established before the start of the experiment using standard procedures, including adaptation, hopper-training, autoshaping, and response-contingent reinforcement, and then through the introduction of a series of variable-interval (VI) schedules of reinforcement. The experiment itself was divided into five phases—baseline, 3.0 mA shock, 4.5 mA shock, 7.0 mA shock and 0.0 mA shock, as follows.

Baseline. Pecking was reinforced according to a two-component chain schedule in which the first component was a fixed-interval 15-sec schedule and the second a variable-interval 30-sec schedule (chain FI 15-sec VI 30-sec). The first component was correlated with the left key light (yellow) and the second with the right key light (red). After the onset of the left yellow key light, the first peck to that key after 15 sec turned off the yellow light and transilluminated the right red key. The first peck to the red key after the VI 30-sec requirement was met then turned off the red key light and operated the feeder for 3 sec. Withdrawal of food was followed by a 1-min blackout. The houselight was illuminated only while a keylight was transilluminated. The baseline condition remained in effect until no trends were evident in pecking in each component for at least five consecutive sessions.

Shock conditions and return to baseline. A stable performance having been established, the procedure was conducted as in the baseline condition except for the addition of a brief (0.03-sec) 3.0 mA shock closely following (after a delay of 0.01 sec) the termination of food delivery. Thus, the last peck in the second component of the chain produced food immediately and, after a delay of just over 3 secs, shock. As in the baseline condition, a 1-min blackout then preceded the onset of the yellow stimulus at the start of the next trial. The blackout served to minimize possible elicitation effects of the shock that might have interfered with responding in the yellow component. After performance returned to approximate baseline levels (which would be expected because of adaptation to the shock), or after the effect of the given shock intensity on responding in each component was roughly stable, shock intensity was increased to 4.5 mA, and then, after either of the same criteria was met again, to 7.0 mA. After the three shock conditions, the baseline condition (0.0 mA shock) was reinstated until an approximate baseline performance was re-established.

Sessions were conducted daily for each subject. Each session ended after 60 complete trials or 120 min, whichever came first, and ranged in duration, from 108 to 120 min.

RESULTS

Figure 1 contains two cumulative record segments showing the typical result of Experiment 1. Record A shows a high, steady, stable rate of responding in red (the VI 30-sec component) and short latencies, variable but low rates of responding, and occasional accelerations in the yellow (FI 15-sec) component.

Record B is from the end of the first session conducted with a 4.5 mA shock for the same subject. Responding in the second component is identical in both rate and temporal pattern with that in Record A, established before shock was added to the schedule. Responding in the first component, however, is dramatically suppressed, with the subject occasionally waiting more than 10 min before a single peck. Similar records were obtained for all three subjects, though the changes were generally less dramatic for one of the subjects (233WP).

This pattern of suppression is reflected in changes in the response rates in each component. The median response rate in the second component for all three subjects and shock conditions was 96% of the average baseline rate, indicating relatively little suppression; the corresponding value for the first component was 28% of the average baseline rate, indicating considerable suppression. Responding in the second component is substantially suppressed in only one of the 26 shock sessions—the first session in which shock was introduced for subject 398WP.

The pattern is also reflected in the average time spent in each component. The top half of Fig. 2 summarizes these data.

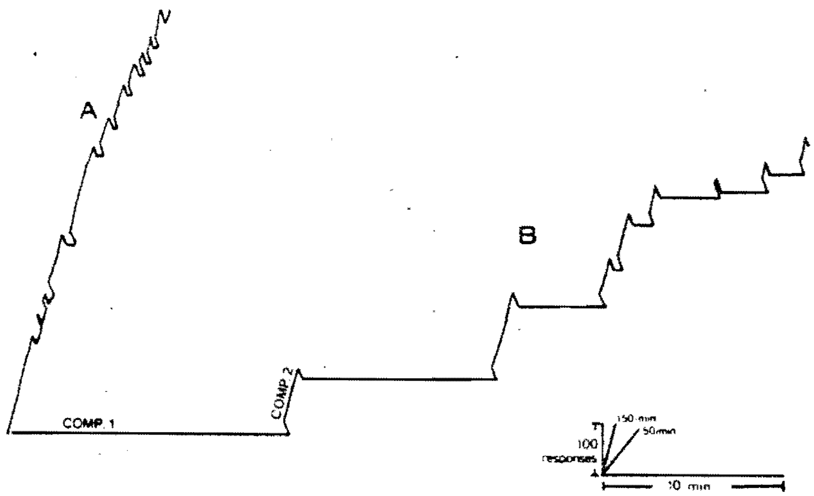

Fig. 1. Two cumulative record segments for Subject 1 (398WP) in Experiment 1. Record A is from the end of the last session conducted without shock in the baseline condition; record B is from the end of the first session during which food was followed by a brief 4.5 mA shock. The pen deflects downward during the first (yellow) component and is in the normal position during the second (red) component. The recorder stops for the subsequent delivery of food and shock and for the 1-min blackout that precedes the onset of the next yellow stimulus. Note that in the shock condition (record B), responding is suppressed in the first component.

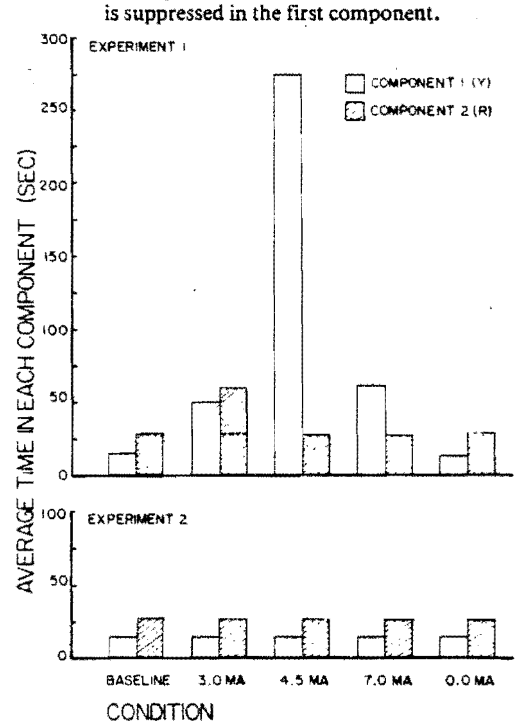

Fig. 2. Average time in each component, averaged across all subjects and sessions, for Experiments 1 and 2. In Experiment 1 (upper graph), a two-component chain (FI 15-sec VI 30-sec) was terminated either with food alone or food followed by shock. The average time spent in the first component increased dramatically during the shock conditions, whereas the average time spent in the second component was affected little. (The increase in time spent in the second component in the 3.0 mA condition was due entirely to the first shock session for subject 398WP. The dotted line indicates the value obtained if this outlying value is not considered.) In Experiment 2 (lower graph), the same schedule was terminated either with food alone or shock followed by food. The shock produced little or no suppression.

DISCUSSION

Several different accounts of these data are possible. One, which is consistent with certain characterizations of the effects of conflicting consequences found in the human self-control literature (e.g. Thoresen and Mahoney, 1974; Kanfer and Phillips, 1970), is as follows. When immediate reinforcement and delayed punishment are imminent, responding is controlled primarily by the reinforcer; only when these consequences are sufficiently delayed is the punishment effective. It might be said that in the second component of the chain a subject could not "resist"; in the first component it could behave more "prudently".

Experiment 1 does not rule out other, simpler accounts of the effect. First, the shock may simply subtract an equal amount of responding from each component. Since there is far less responding in the first component to begin with, we see little suppression in the second component and a great deal in the first. Second, the effect may be due to elicitation. It is possible that at the end of the 1-min blackout the bird has still not recovered from the last shock. When it finally recovers, it begins to peck as usual. The first peck in component 1 satisfies the FI 15-sec requirement, and the subsequent rate and pattern of responding in component 2 resemble those of the baseline sessions. Third, the effect may be an artifact of the schedules employed. Rates on VI schedules tend to get "locked-in"; FI responding may simply be more sensitive to the addition of shock. Fourth, early components in chain schedules of reinforcement are known to be more disruptible than later ones (Nevin, 1979; cf. Myer, 1973).

Experiments 2 and 3 below control for these possibilities. The subtraction, elicitation and disruptibility accounts may be tested (though not distinguished) by presenting the shock before reinforcement instead of after it. Presenting the shock 3 sec sooner should make little difference if elicitation is responsible for the effect, and should enhance the effect if shock is subtracting responses from each component or if the effect is due to the relative disruptibility of responding in the first component. In Experiment 2, shock is presented immediately before food and in Experiment 3, the effect of presenting shock before food is compared with the effect of presenting it after food in a within-subjects design. A possible schedule interaction is also tested in Experiment 3.

EXPERIMENT 2

METHOD

Subjects and apparatus

Three experimentally naive, adult male White Carneaux pigeons, numbers 403WP, 401WP and 276WP, served as subjects. Each was maintained at about 80% of its free-feeding weight and was wired for shock as in Experiment 1. The apparatus was the same as in Experiment 1.

Procedure

Baseline performances were established as in Experiment 1, using the same schedule of reinforcement (chain FI 15-sec VI 30-sec, followed by 3 sec of food and a 1-min blackout), after which a brief (0.03-sec) shock was inserted into the sequence just before the onset of reinforcement (the feeder was operated 0.01 sec after the termination of shock). Shock intensity was increased from 3.0 to 7.0 mA over eight sessions for subject 403WP and over five sessions for subjects 401WP and 276WP, after which baseline conditions were re-established, as in Experiment 1.

RESULTS

Inserting the shock before the food resulted in little suppression in either component, at any shock value. The average time spent in each component, averaged across all subjects and sessions, is pictured in the lower half of Fig. 2. The dramatic increases that occurred when shock came after food (top graph) did not occur when shock came before food.

DISCUSSION

Since the suppression pattern that was found in Experiment 1 did not occur again in Experiment 2, it probably was not due to either subtraction, elicitation, or the relative disruptibility of responding in the first component. That it was due to the schedules also seems unlikely, unless the pattern was determined by both the schedules and the position of the shock. Experiment 3 uses a within-subjects design to test these possibilities.

EXPERIMENT 3
METHOD

Four experimentally naive, adult male White Carneaux pigeons, numbers 289WP, 291WP, 406WP and 411WP served as subjects. Each was maintained at about 80% of its free-feeding weight and was wired for shock as in Experiment 1. The apparatus was the same as in Experiment 1.

Procedure

In the baseline condition, pecks were reinforced according to a chain FI 15-sec FI 15-sec schedule. As in the previous experiments, the first component was correlated with a yellow light on the left key and the second component with a red light on the right key. Satisfying the second FI requirement always produced 3 sec of food followed by a 1-min blackout. After a stable baseline performance was established, two of the subjects (289WP and 291WP) received, for two sessions, a brief (0.03-sec) 3.0 mA shock immediately after the termination of food, as in Experiment 1. Baseline conditions were then reinstated for seven sessions, after which a shock of the same intensity and duration was inserted immediately before food for two sessions, as in Experiment 2. The order of the two shock conditions was reversed for subjects 406WP and 411WP. Daily sessions were always 1 hr in length.

RESULTS

The results of Experiments 1 and 2 were confirmed. Responding in the second component was affected little by shock placed either before or after food. The subjects (289WP and 291WP) that first received food followed by shock produced the suppression pattern that was observed in Experiment 1. There was substantial suppression only in the first component. When they later received shock before food, there was little suppression in either component. The subjects (406WP and 411WP) that first received shock before food showed little or no suppression in either component. When they later received food followed by shock, the fourth subject (411WP) showed some first-component suppression; no suppression was evident in either component for the third subject (406WP). (A small or no effect of shock during the second exposure is presumably the result of adaptation. This possibility applies to all four of the subjects.)

The result is summarized in Fig. 3, which shows the average time spent in each component, averaged across all subjects and sessions. The average time spent in the second component is between 15.4 and 15.9 sec for shock and baseline conditions alike. The average time spent in the first component is 111 sec when shock follows food, but is otherwise between 17.3 and 19.2 sec.

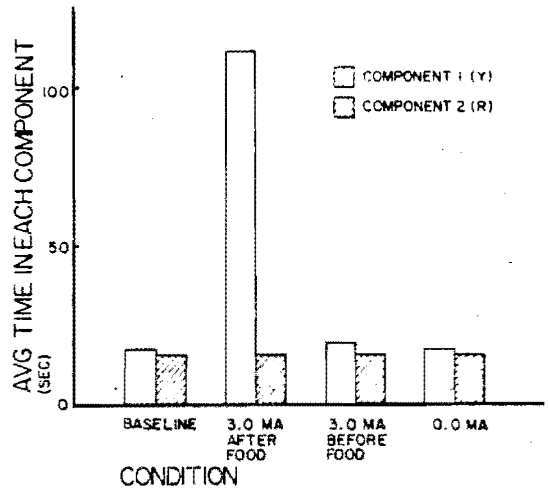

Fig. 3. Average time spent in each component, averaged across all subjects and sessions for Experiment 3. A two-component chain (FI 15-sec FI 15-sec) was terminated either with food alone, with food followed by shock, or with shock followed by food. The average time spent in the first component increased dramatically when food was followed by shock. The average time spent in the second component was unaffected by the shock in either shock condition.

GENERAL DISCUSSION

The pattern of suppression observed in Experiments 1 and 3 when shock followed food is harmonious with both casual and professional observations about the effects of immediate reinforcement and delayed punishment on human behavior. The overeater, faced with the chocolate cake, is not likely to resist; the smoker, even the one trying to quit, is likely to smoke if cigarettes are at hand. Faced with immediate reinforcement and delayed punishment—in other words, "temptation"—the reinforcer is typically the effective stimulus. If, on the other hand, the reinforcer is somehow less immediate, behavior related to the delayed punishment may prevail. (Fig. 4). As long as the cake is out of sight, the overeater may adhere to his or her diet; while not around smokers, the person trying to quit may make plans to be with non-smokers.

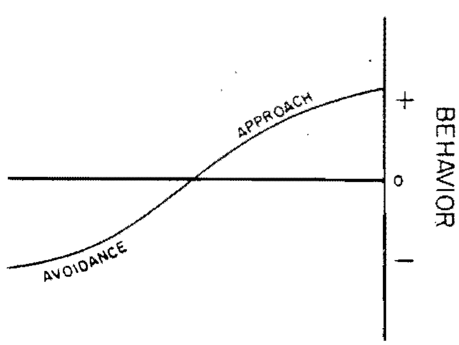

Fig. 4. An approach-avoidance model of self-control. When a reinforcer which is followed by delayed punishment (right-hand portion of the curve) is imminent, approach behavior is likely. When such a reinforcer is more remote in time (left-hand portion), avoidance is likely. A simple recency account is possible.

Reinforcers often accompanied by delayed punishment are prevalent and so problematic in their effects that they have for many years been a major concern of therapists and are central to several conceptions of human "self-control" (e.g. Skinner, 1953; Ferster, Nurnberger and Levitt, 1962; Thoresen and Mahoney, 1974; Kanfer and Phillips, 1970). Self-control behavior is said to be responsive to "ultimate aversive consequences" (Ferster, 1965), to "'bridge' the gradient" between immediate and delayed consequences (Thoresen and Mahoney, 1974), and to remove "potentially maladapted sources of gratification" (Goldfried and Merbaum, 1973); which is to say, it is often behavior that somehow avoids a reinforcer that has previously been followed by delayed punishment.

The key to self-control may be the acquisition of a large number of avoidance behaviors relevant to reinforcers correlated with delayed punishment. The avoidance behaviors may take any number of forms: inaction ("abstinence"), alternative or incompatible behavior, pre-commitment (cf. Epstein and Goss, 1978; Rachlin, 1974), and so on.

The suppression pattern observed in Experiments 1 and 3 may also be interpreted in terms of the choice-based model of self-control currently the concern of many researchers (e.g. Ainslie, 1974, 1975; Rachlin, 1970, 1974; Rachlin and Green, 1972; Mazur and Logue, 1978; Deluty, 1978; Navarick and Fantino, 1976; Green and Snyderman, 1980; Solnick, Kannenberg, Eckerman and Waller, 1980), though the applicability of the model is not clear. Abstaining from pecking in the early component could be considered foregoing a near, less "valued" reinforcer—food followed by shock—for some more remote, more "valued" reinforcer—perhaps the food given to the bird after it is removed from the chamber to maintain it at running weight. Shock followed by food must somehow be more "valuable" than food followed by shock, according to this model, since it does not produce this suppression pattern. The choice might also be said to be between immediate food and the omission of delayed shock, but this interpretation is lacking in two respects. First, the first alternative involves *both* reinforcement and (delayed) punishment; a response cannot produce one without producing the other, as is often the case for human behavior. Second, the effect of the second alternative should be greater when shock comes sooner; thus, again, shock before food should produce greater suppression than shock after food, which it does not.

Neither the conflicting consequences paradigm nor the choice model makes contact with important elements of human self-control, as characterized by both professionals and laypersons. Self-control is often characterized as a "decision" process, not a choice process (Kanfer, 1977). Unlike the passive and inevitable behavior of a pigeon that occurs as a function of the remoteness of various events, it is said to involve some "deliberate" action, some manipulation of behavior or environment related to temporally remote factors; hence the force of the term "control" in "self-control" and the word's kinship with "self-denial", "willpower" and so on (cf. Bandura, Mahoney and Dirks, 1976; Ferster, 1965; Mahoney and Bandura, 1972; Skinner, 1953). Dictionary definitions emphasize "self-denial" and "self-restraint"—in other words, resisting temptation even when it is sufficiently near that its effectiveness curve is presumably still higher than that of some long-term consequence. "Self-control" is said to refer to behavior that young children (and presumably pigeons) lack entirely (Bandura and Walters, 1963), that is acquired only after an extensive history of training and socialization (Premack and Anglin, 1973; Bandura and Walters, 1963; Risley, 1977; cf. Mazur and Logue, 1978), and that many people never acquire to any great extent. It seems unlikely that pigeons with no special training history are capable of behavior that takes humans years to learn, and imperfectly at that.

Moreover, most instances of behavior that come under the rubric of self-control are the result neither of conflicting consequences nor of remote alternatives that one has actually experienced. They are often simply imitated or under instructional control. Odysseus did not experience the deleterious consequences of the Sirens' song before having himself tied to his mast (rather sophisticated avoidance behavior); he simply followed Circe's instructions (cf. Ainslie, 1975; Rachlin and Green, 1972). And one need not have experienced the immediate reinforcement and delayed punishment of oversleeping to set an alarm clock nor have contracted cancer to stop smoking. Nevertheless, to the extent that remote, unequal alternatives or the effects of conflicting consequences on a single response are involved in self-control phenomena, current laboratory procedures may shed light on them.

REFERENCES

Ainslie G. (1974) Impulse control in pigeons. *J. Exp. Analysis Behav.* **21**, 485–489.

Ainslie G. (1975) Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychol. Bull.* **82**, 463–496.

Azrin N. H. (1959) A technique for delivering shock to pigeons. *J. Exp. Analysis Behav.* **2**, 161–163.

Azrin N. H. and Holz W. C. (1966) Punishment. In *Operant Behavior: Areas of Research and Application* (Ed. by Honig W. K.). Prentice-Hall, Englewood Cliffs, New Jersey.

Bandura A., Mahoney M. J. and Dirks S. J. (1976) Discriminative activation and maintenance of contingent self-reinforcement. *Behav. Res. Ther.* **14**, 1–6.

Bandura A. and Walters R. (1963) *Social Learning and Personality Development.* Holt, Rinehart & Winston, New York.

Deluty M. (1978) Self-control and impulsiveness involving aversive events. *J. Exp. Psychol.: Anim. Behav. Proc.* **4**, 250–266.

~~Dmitruk V. M. (1974) 'Experimental neurosis' in cats: Fact or artifact? *J. abnorm. Psychol.* **83**, 97–105.~~

Epstein R. and Goss C. M. (1978) A self-control procedure for the maintenance of nondisruptive behavior in an elementary school child. *Behav. Ther.* **9**, 109–117.

Ferster C. B. (1965) Classification of behavioral pathology. In *Research in Behavior Modification* (Ed. by Krasner L. and Ullmann L.). Holt, Rinehart & Winston, New York.

Ferster C. B., Nurnberger J. I. and Levitt E. B. (1962) The control of eating. *J. Mathetics* **1**, 87–109.

Gantt W. H. (1944) Experimental basis for neurotic behavior. *Psychosom. Med. Monogr.* **3**, Nos. 3 and 4.

Goldfried M. and Merbaum M. (1973) A perspective on self-control. In *Behavior Change Through Self-control* (Ed. by Goldfried M. and Merbaum M.). Holt, Rinehart & Winston, New York.

Green L. and Snyderman M. (1980) Choice between rewards differing in amount and delay: Toward a choice model of self-control. *J. Exp. Analysis Behav.* **34**, 135–147.

Holz W. C. (1968) Punishment and rate of positive reinforcement. *J. Exp. Analysis Behav.* **11**, 285–292.

Kanfer, F. (1977) The many faces of self-control, or behavior modification changes its focus. In *Behavioral Self-management: Strategies, Techniques, and Outcomes* (Ed. by Stuart R.). Brunner/Mazel, New York.

Kanfer F. H. and Phillips J. S. (1970) *Learning Foundations of Behavior Therapy.* Wiley, New York.

Mahoney M. J. and Bandura A. (1972) Self-reinforcement in pigeons. *Learn. Motiv.* **3**, 293–303.

Masserman J. H. (1943) *Behavior and Neurosis: An Experimental Psychoanalytic Approach to Psychobiologic Principles.* Chicago University Press, Chicago.

Mazur J. and Logue A. (1978) Choice in a 'self-control' paradigm: Effects of a fading procedure. *J. Exp. Analysis Behav.* **30**, 11–17.

Miller N. E. (1937) Reaction formation in rats: An experimental analog for a Freudian phenomenon. *Psychol. Bull.* **34**, 724.

Miller N. E. (1959) Liberalization of basic S–R concepts: Extensions to conflict behavior, motivation, and social learning. In *Psychology: A Study of a Science,* Vol. 2 (Ed. by Koch S.). McGraw-Hill, New York.

Myer J. S. (1973) Effects of punishing elements of a simple instrumental-consummatory response chain. *J. exp. Analysis Behav.* **19**, 251–257.

Navarick D. and Fantino E. (1976) Self-control and general models of choice. *J. exp. Psychol.: Anim. Behav. Proc.* **2**, 75–87.

Nevin J. A. (1979) Reinforcement schedules and response strength. In *Reinforcement and the Organization of Behavior* (Ed. by Zeiler D. and Harzem P.). John Wiley, Chichester.

Premack D. and Anglin B. (1973) On the possibilities of self-control in man and animals. *J. abnorm. Psychol.* **81**, 137–151.

Rachlin H. (1970) *Introduction to Modern Behaviorism.* Freeman, San Francisco.

Rachlin H. (1974) Self-control. *Behaviorism* **2**, 94–107.

Rachlin H. and Green L. 1972. Commitment, choice, and self-control. *J. Exp. Analysis Behav.* **17**, 15–22.

Risley T. R. (1977) The social context of self-control. In *Behavioral Self-management: Strategies, Techniques, and Outcomes* (Ed. by Stuart R.). Brunner/Mazel, New York.

Skinner B. F. (1953) *Science and Human Behavior.* Macmillan, New York.

Solnick J., Kannenberg C. H., Eckerman D. A. and Waller M. B. (1980) An experimental analysis of impulsivity and impulse control in humans. *Learn. Motiv.* **11**, 61–77.

Stunkard A. (1977) Behavioral treatment of obesity: Failure to maintain weight loss. In *Behavioral Self-management: Strategies, Techniques, and Outcomes* (Ed. by Stuart R.). Brunner/Mazel, New York.

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

Acknowledgements—The author is grateful to R. J. Herrnstein, J. E. Mazur, B. F. Skinner and W. Vaughan, Jr. for support and suggestions.