

Amsel's analysis of reward-schedule effects

M. E. BITTERMAN

University of Hawaii, Honolulu, Hawaii

Some questions raised by Amsel's valuable review of his extensive research on reward-schedule effects are considered—questions about the nature of instrumental conditioning, about the distinction between long- and short-term memory, and about evolutionary divergence in learning.

We are greatly indebted to Amsel, not only for his long and important work on frustration and learning, but now also for his comprehensive review of that work (Amsel, 1992), and I am glad of the opportunity to comment here on three matters of special interest to me which it brings to mind. One valuable contribution of the review is that it prompts us to reconsider a powerful approach to the problem of instrumental conditioning which has been too long in eclipse; another is that it calls into sharp focus the question of the relation between short- and long-term memory; and a third is that it dramatizes the limitations of contemporary thinking about the evolution of learning.

Instrumental Conditioning

In this time of reawakened anthropomorphism, with Morgan's canon daily being set on its ear, Amsel's unabashedly Hullian analysis, expressed in terms of what many now find to be tiresome Ss and Rs, may seem hardly to be taken seriously, but I see Hullian theory, for all its limitations, as our most promising starting point on the road to something better. Although the S-R reinforcement principle—the centerpiece of the theory—is commonly supposed to have been contradicted almost from the outset by evidence from experiments on incentive contrast that the consequences of response are in some sense learned “about” (Tolman, 1932), the proper conclusion from the evidence is only that the principle is insufficient, and no one has ever been able to show convincingly how we can get along without it.

When an animal makes some response to a stimulus and is rewarded with food, the stimulus may be associated with the food—a prospect long recognized (Morgan, 1894)—but there is no way to account in terms of stimulus-reinforcer association for the fact that the instrumental response can be different from the response evoked by the stimulus when it is paired with the food independently of the animal's behavior (that is, in a purely classical procedure). Instrumental performance is depressed after the reward has been devalued (paired with

illness), but usually only incompletely, especially after prolonged training; that is, the animal may continue to work for a food that it refuses (e.g., Adams, 1980), which, of course, is what the S-R reinforcement principle leads us to expect.

The equally old (Morgan, 1894) and equally reasonable assumption that interoceptive feedback from the instrumental response may be associated with the food (the traditional meaning of “response-reinforcer” association) does not explain why the rewarded response is increasingly likely to be elicited, nor does the more recent assumption that the stimulus is associated with an *association* between the response and the food (Rescorla, 1990). Dissatisfaction with the S-R reinforcement principle has led also to some unconvincing proposals to substitute innate S-R associations for acquired ones. Mowrer (1956) was willing to assume, for example, that the animal constantly scans all the responses in its repertoire, sampling the feedback from each, until it hits upon one whose feedback has in the past been associated with reward. At a time when cybernetics was in fashion, I invented that absurd possibility simply to make a point in a discussion with Mowrer, who surprised me by taking it seriously; and there have been other surprises.

After a thoughtful review of a variety of efforts to do without the S-R reinforcement principle, none of which he found acceptable, Mackintosh (1983) offered a neo-Tolmanian alternative: “The association between instrumental response and reinforcer should be regarded,” he wrote,

as an expectation or proposition about certain relationships, which, when combined with other premises, can be used to derive instructions to perform or withhold that response.... Instrumental performance is thus the consequence of an instruction inferred from an association, rather than being directly elicited by an association.” (p. 111)

In the same vein, Dickinson (1989) has recommended Brentano's

concept of intentionality ... to the property of mental states to possess a content.... This view of knowledge about the world allows us to conceive of the current value of a reinforcer as a form of practical inference [that] takes the content of the belief about the conditioning contingency and the content of a desire for the reinforcer as

Correspondence should be addressed to M. E. Bitterman, Békésy Laboratory of Neurobiology, 1993 East-West Road, Honolulu, HI 96822 (e-mail: jeffb@ahi.pbrc.hawaii.edu).

premises and yields as a conclusion a command to perform the conditioned response or action." (p. 284)

In retrospect, those old Ss and Rs may not seem quite so objectionable.

If we do adopt the S-R reinforcement principle to explain instrumental conditioning, but not the clearly untenable early suggestion that classical conditioning also can be explained in terms of that principle (Hull, 1943), we return, of course, to two-process theory and to the old problem of distinguishing the two processes. The S-R contiguity account of classical conditioning that Amsel prefers is awkward in several respects: Reward is assumed to be anticipated on the basis, not of its immediate sensory properties, but of feedback from the consummatory response to it; for different rewards to be distinguishable, they must evoke either different consummatory responses or the same response with different vigor (cf. Spence, 1956). Sensory preconditioning—which suggests that contiguous *stimuli* are associated (Maier & Schneirla, 1942)—can be explained in terms of the S-R contiguity principle only on the dubious assumption that each of the paired neutral stimuli is connected in the first stage of training to a subtle yet distinctive response evoked by the other, the feedback from which is connected to the observable target response in the second stage (cf. Bitterman, Reed, & Kubala, 1953).

The assumption that classical conditioning is an S-R process implies, furthermore, that there are two categories of conditionable responses, one requiring reinforcement and the other not. Of the various dichotomies that have been proposed, the only unambiguous one—that between autonomic responses said to be classically conditionable and skeletal responses said to be instrumentally conditionable—is invalidated by the classical conditioning of leg flexion in dogs, eyelid closure in rabbits, and keypecking in pigeons and goldfish, which also can be conditioned instrumentally (although there is no good evidence that autonomic responses can be instrumentally conditioned). Another old dichotomy—that between "voluntary" and "involuntary" responses—has yet to be given any operational meaning, at least in work with animals. The traditional behaviorist reluctance to consider the S-S contiguity account of classical conditioning, which faces none of these difficulties, seems based on the unwarranted impression that hypothetical S-S associations are somehow less objective than equally hypothetical S-R associations.

Whatever the nature of the classical process, just how it influences instrumental performance is far from clear. Amsel assumes that feedback from a classically conditioned response that antedates the rewarded instrumental response is associated with it, leading, of course, to his convincing explanation of the partial reinforcement extinction effect in terms of counterconditioning—the partial animals are said to be rewarded in training for response to feedback from conditioned frustration. Classical conditioning is treated also as a source of appetitive motivation. There is evidence, however, that expectation

of food is evoked, not by the CS (which is never alone paired with the food), but by the CS in compound with feedback from the instrumental response (the response seeming to trigger the expectation rather than the other way round), and that a stimulus paired independently with food may actually interfere with instrumental responding (Konorski, 1967). We have some fascinating questions here, long obscured in clouds of cognitive metaphor, to which Amsel's analysis invites our renewed attention.

Long- and Short-Term Memory

At the outset of his review, Amsel distinguishes between "dispositional" or long-term memories (the stable products of associative learning that have been of primary interest to him) and short-term or working memories (rapidly decaying residuals of recent stimulation). Short-term memories may themselves, of course, acquire (long-term) associative control of responding. In delayed "symbolic" matching-to-sample, for example, a recently seen square may signal to a pigeon that choice of red rather than green will be rewarded, and a recently seen circle may signal that choice of green rather than red will be rewarded; if the pigeon is to perform successfully, the interval between the presentation of the sample and the rewarded choice must be very short, although the interval between trials may be long. The distinction between long- and short-term memory—neither conception being necessarily more or less objective (or even Hullian) than the other—is a simple one that can be made easily without reference to Squire's fuzzy table of descriptors.

The theory that reward-schedule effects in massed trials may reflect associative control of performance by sensory aftereffects or short-term memories of the consequences of response carried over from one trial to the next is not, I think, given its just due by Amsel, who sometimes seems still to treat it as a competitor rather than a companion of frustration theory. The carryover interpretation certainly is not *contradicted*, as Amsel implies it is, by the fact that resistance to extinction may be greater after partial than after consistent reinforcement even when extinction is preceded in both cases by a series of consistently reinforced trials. An animal that has been rewarded in the past for responding to aftereffects of nonreward should be expected to respond again to those aftereffects when they recur following a period of consistent reward if the association between the aftereffects of nonreward and the instrumental response is itself "dispositional."

Reward-schedule effects that occur in spaced trials are clearly beyond the reach of the carryover interpretation, but there is a good deal of evidence that carryover does play a broad and important role in massed trials. Trained with small reward, rats show the partial reinforcement (extinction) effect in massed trials but not in spaced trials, where presumably there is neither carryover nor appreciable frustration (e.g., Gonzalez & Bitterman, 1969). Fishes of divergent species—African

mouthbreeders and goldfish—show the partial reinforcement effect in massed trials but never at all in spaced trials, even with large reward (Longo & Bitterman, 1960; Schutz & Bitterman, 1969). Although the effect is less likely to appear in classical than in instrumental conditioning, it has been found in the classical conditioning of rats and African mouthbreeders with shock as the unconditioned stimulus (Crum, Brown, & Bitterman, 1951; Gonzalez, Eskin, & Bitterman, 1963), an outcome more readily understandable in terms of carryover than of conditioned frustration. The patterned responding that develops in rats and goldfish (Couvillon & Bitterman, 1981; Tyler, Wortz, & Bitterman, 1953) when rewarded and nonrewarded trials are singly alternated—a schedule effect characterized by Amsel (1992, p. 181) as more “primitive” than the others and inexplicable in terms of conditioned frustration—points clearly to control by carryover.

Is the conditionable carryover sensory or mnemonic? If the two can be distinguished to some extent in terms of susceptibility to interference by interpolated stimuli, there is evidence of both. In one of a series of experiments by Couvillon, Brandon, Woodard, and Bitterman (1980), pigeons trained in massed trials with response to a red key (S_3) rewarded and nonrewarded in single alternation came to respond rapidly on rewarded trials and slowly on nonrewarded trials, but the patterning disappeared immediately when trials with two other colors—response to one ($S+$) consistently rewarded and to the other ($S-$) never—were interpolated between successive trials with S_3 . An interpretation of the patterning based on the discrimination of sensory carryover is supported by the fact that the pigeons continued for many sessions to respond rapidly to S_3 after trials with $S-$ and slowly after trials with $S+$.

Evidence of associative control by short-term memory of nonreward was found in a companion experiment on successive acquisition and extinction of response to S_3 , again with interpolated $S+$ and $S-$ trials (Couvillon

et al., 1980). One group of pigeons was trained with S_3 positive in two sessions, negative in the next two sessions, and so forth; that is, the reversals of sign were between sessions. For a second group of pigeons, the training was the same except that the reversals were within sessions; at the midpoint of every other session, S_3 became positive after having been negative, or negative after having been positive. Plotted in Figure 1 is the asymptotic performance of both groups in sessions with S_3 consistently negative. The striking difference in resistance to extinction is explained on the assumption that response to short-term memory of S_3 negative is rewarded in within-sessions reversal training (where S_3 is first negative and then positive in the same session) but not in between-sessions reversal training. It is difficult to understand how these results might be accounted for in terms of long-term memory alone.

The results of an analogous reversal experiment by Couvillon (1984) with goldfish (which show the partial reinforcement effect in massed but not in spaced trials) also are plotted in Figure 1; the similar within-sessions and between-sessions curves give no indication that the performance of the goldfish was controlled by short-term memory of nonreward. The conventional partial reinforcement effect was sought directly in both the pigeon and the goldfish experiments, again under conditions in which interpolated trials with other stimuli were used to minimize sensory carryover. As indicated by the extinction curves plotted in Figure 2, the pigeons showed the effect but the goldfish did not. Both comparisons suggest that goldfish must rely entirely on sensory carryover for guidance in reward-schedule experiments.

As to long-term association (“dispositional memory”) itself, I am puzzled by Amsel’s assumption that it is permanent—that the influence of a long-term association (however temporally stable) can be erased only by counterconditioning. In Hullian theory, both $S_H R$ and $S_I R$ were assumed to be permanent, which made it difficult

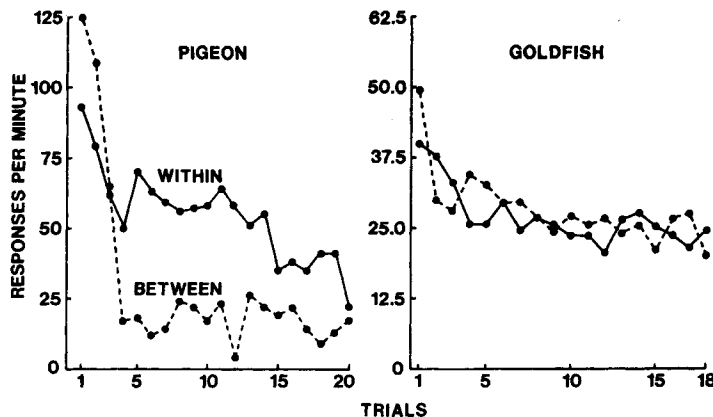


Figure 1. Asymptotic resistance to extinction of pigeons and goldfish given successive acquisition-extinction training in which the transitions from reward to nonreward and from nonreward to reward occurred either within or between sessions. Interpolated trials with other stimuli were used to control for sensory carryover.

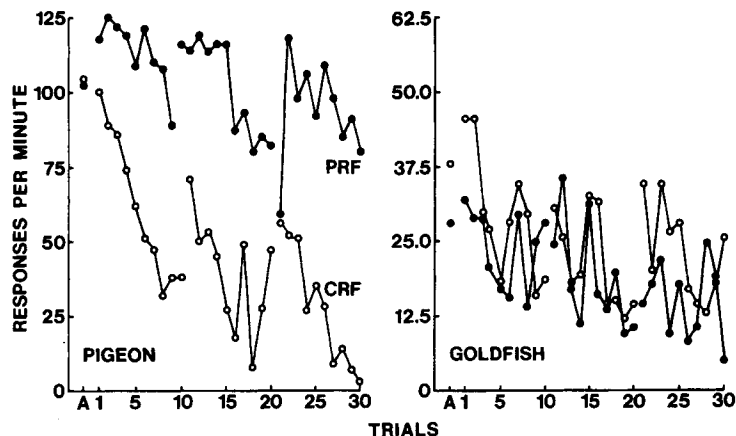


Figure 2. Resistance to extinction in pigeons and goldfish in three sessions following partial reinforcement (PRF) or consistent reinforcement (CRF) in experiments in which interpolated trials with other stimuli were used to control for sensory carryover. Mean rate of responding in the preceding training session is plotted at A.

to account for sustained flexibility in the repeated conditioning and extinction of a single stimulus, or in discrimination training with repeated reversals of the positive and negative stimuli (cf. Gleitman, Nachmias, & Neisser, 1954), and the addition of frustration as another permanent factor (or even as a replacement for S_{IR}) does not make things any better. In massed trials, as Hull (1952) noted, carryover may be called upon—how successfully remains to be determined—but, in spaced trials at least, some sort of reduction in associative strength is required if permanent inflexibility is not to ensue as the competing associations approach their asymptotes (Woodard & Bitterman, 1976).

Consider the simpler question of why the successive negative contrast effect found in spaced trials is transient—why the performance of the group shifted from preferred to less preferred reward, although disrupted at first, comes after a time to equal that of the unshifted control group: Even if the original expectation of the preferred reward is replaced by an expectation of the less preferred reward (counterconditioning?) and primary frustration is in consequence no longer generated, how (without extinction of conditioned frustration) can the improvement in performance be explained? It is interesting to note that in their attempt at a quantitative statement of frustration theory, Daly and Daly (1982) have permitted associative strength not only to decline, but even to become negative. The conception of permanent association is one Hullian legacy that does not seem worth preserving.

Evolutionary Divergence in Learning

The early theory of Thorndike (1911) that the laws of learning are the same in all animals was widely accepted for many years and then abandoned precipitously under the mysterious influence of ethologists who knew essentially nothing about learning. A popular opinion today is that there may indeed be some common mech-

anisms of learning, but that they have undergone a good deal of species-specific evolutionary tailoring designed to meet varying ecological demands (e.g., Gould & Marler, 1984), and speculation about the adaptive advantage conferred by this or that supposed specialization in the learning of one or another animal is a favorite contemporary occupation. In some quarters, complacent speculation about adaptive significance has come almost to replace functional analysis. The proponents of specialization seem not to appreciate what is required to demonstrate a difference in the learning of two species. Learning is not, after all, given directly in behavior, but inferred from behavior that is determined also by many other variables—motivation, for example—whose effects must somehow be controlled before conclusions about the contribution of learning are warranted (Bitterman, 1975). In my view, there is as yet no evidence of specialization that will stand up to critical scrutiny (Bitterman, 1976; Bitterman & Couvillon, 1991).

An intriguing indication of a broader evolutionary divergence in the learning of vertebrates does come, however, from comparative work on reward-schedule effects (Bitterman, 1987). The phenomenon of successive negative incentive contrast that first pointed conclusively to the insufficiency of the S-R reinforcement principle for rats and monkeys has failed repeatedly to appear in animals of older vertebrate lines (fishes, toads, turtles). So also have the inverse relation between amount of reward and resistance to extinction (which can be treated as a special case of negative contrast) and the *spaced-trials* partial reinforcement effect (of which precipitous extinction after consistent large reward is a critical feature). Failures to find the three effects may mean only that they have not been sought in the right places, although that is increasingly unlikely in view of the rather wide variety of experimental conditions employed and the evolutionary relation of species that do and do not show the effects (Bitterman, 1975). Given that the missing effects

are related by frustration theory, it has been tempting to think in terms of a set of underlying mechanisms, not general to vertebrates, but appearing for the first time in some reptilian ancestor of birds and mammals.

The hypothesis of divergence is considerably strengthened by Amsel's beautiful ontogenetic experiments with rats, which provide further evidence (in the youngest subjects) of what can be conceived of as pure or non-paradoxical instrumental learning—learning that does not challenge the sufficiency of the S-R reinforcement principle. As Amsel has recognized, the methodological problems are no less difficult in ontogenetic comparisons than in phyletic comparisons for the reason again that performance is influenced by variables other than learning, but the ontogenetic comparisons like the phyletic comparisons have been made under a relatively wide range of conditions, and the two sets of results draw support from each other.

A special difficulty inherent in the work with very young rats is that the training is massed, which may be the reason for the relatively early appearance of the partial reinforcement effect. It may be well to note that the ontogenetic results go further than the phyletic results in emphasizing that we are dealing (apart from the carry-over question) with a range of paradoxical effects that do not necessarily appear, or fail to appear, together. In massed postshift trials, for example, successive negative contrast can be accounted for in terms of unconditioned frustration alone (without reference either to conditioned or to counterconditioned frustration) and, in spaced postshift trials, in terms both of conditioned and unconditioned frustration (without reference to counterconditioning), while the partial reinforcement effect (absent carryover) requires all three. The successive negative contrast found in honeybees (Couvillon & Bitterman, 1984) can be accounted for in terms of unconditioned frustration alone, as can the *simultaneous* negative contrast found in goldfish (Burns, Woodard, Henderson, & Bitterman, 1974). The ontogenetic results point to the need for a more analytical approach in the phyletic work, which has thus far been directed primarily to the question of whether there has been any evolutionary divergence in learning at all. Amsel's ontogenetic strategy (outlined in Table 3 of Amsel's [1994] précis) recommends itself also as a phyletic strategy and one whose usefulness extends beyond the study of reward-schedule effects to the study of behavior in general.

REFERENCES

- ADAMS, C. D. (1980). Post-conditioning devaluation of an instrumental reinforcer has no effect on extinction performance. *Quarterly Journal of Experimental Psychology*, **32**, 447-458.
- AMSEL, A. (1992). *Frustration theory: An analysis of dispositional learning and memory*. New York: Cambridge University Press.
- AMSEL, A. (1994). Précis of *Frustration theory: An analysis of dispositional learning and memory*. *Psychonomic Bulletin & Review*, **1**, 280-296.
- BITTERMAN, M. E. (1975). The comparative analysis of learning. *Science*, **188**, 699-709.
- BITTERMAN, M. E. (1976). Issues in the comparative psychology of learning. In R. B. Masterton, M. E. Bitterman, C. B. G. Campbell, & N. Hottot (Eds.), *Evolution of brain and behavior in vertebrates* (pp. 217-225). Hillsdale, NJ: Erlbaum.
- BITTERMAN, M. E. (1987). Evidence of divergence in vertebrate learning. *Behavioral & Brain Sciences*, **10**, 659-660.
- BITTERMAN, M. E., & COUVILLON, P. A. (1991). Failures to find evidence of adaptive specialization in the learning of honeybees. In L. J. Goodman & R. C. Fischer (Eds.), *The behaviour and physiology of bees* (pp. 188-305). Wallingford, U.K.: CAB International.
- BITTERMAN, M. E., REED, P. C., & KUBALA, A. L. (1953). The strength of sensory preconditioning. *Journal of Experimental Psychology*, **46**, 178-182.
- BURNS, R. A., WOODARD, W. T., HENDERSON, T. B., & BITTERMAN, M. E. (1974). Simultaneous contrast in the goldfish. *Animal Learning & Behavior*, **2**, 97-100.
- COUVILLON, P. A. (1984). Performance of goldfish (*Carassius auratus*) in patterned sequences of rewarded and nonrewarded trials. *Journal of Comparative Psychology*, **98**, 333-344.
- COUVILLON, P. A., & BITTERMAN, M. E. (1981). Analysis of alternation patterning in goldfish. *Animal Learning & Behavior*, **9**, 169-172.
- COUVILLON, P. A., & BITTERMAN, M. E. (1984). The overlearning-extinction effect and successive negative contrast in honeybees (*Apis mellifera*). *Journal of Comparative Psychology*, **98**, 100-109.
- COUVILLON, P. A., BRANDON, S. E., WOODARD, W. T., & BITTERMAN, M. E. (1980). Performance of pigeons in patterned sequences of rewarded and nonrewarded trials. *Journal of Experimental Psychology: Animal Behavior Processes*, **2**, 137-154.
- CRUM, J., BROWN, L. W., & BITTERMAN, M. E. (1951). The effect of partial and delayed reinforcement on resistance to extinction. *American Journal of Psychology*, **64**, 228-237.
- DALY, H. B., & DALY, J. T. (1982). A mathematical model of reward and aversive nonreward: Its application in over 30 appetitive learning situations. *Journal of Experimental Psychology: General*, **111**, 441-480.
- DICKINSON, A. (1989). Expectancy theory in animal conditioning. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 278-308). Hillsdale, NJ: Erlbaum.
- GLEITMAN, H., NACHMIAS, J., & NEISSER, U. (1954). The S-R reinforcement theory of extinction. *Psychological Review*, **61**, 23-33.
- GONZALEZ, R. C., & BITTERMAN, M. E. (1969). Spaced trials partial reinforcement effect as a function of contrast. *Journal of Comparative & Physiological Psychology*, **67**, 94-103.
- GONZALEZ, R. C., ESKIN, R. M., & BITTERMAN, M. E. (1963). Further experiments on partial reinforcement in the fish. *American Journal of Psychology*, **76**, 366-375.
- GOULD, J. L., & MARLER, P. (1984). Ethology and the natural history of learning. In P. Marler & H. S. Terrace (Eds.), *The biology of learning* (pp. 47-74). Berlin: Springer-Verlag.
- HULL, C. L. (1943). *Principles of behavior*. New York: Appleton-Century.
- HULL, C. L. (1952). *A behavior system*. New Haven, CT: Yale University Press.
- KONORSKI, J. (1967). *Integrative activity of the brain*. Chicago: University of Chicago Press.
- LONGO, N., & BITTERMAN, M. E. (1960). The effect of partial reinforcement with spaced practice on resistance to extinction in the fish. *Journal of Comparative & Physiological Psychology*, **53**, 169-172.
- MACKINTOSH, N. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press, Clarendon Press.
- MAIER, N. R. F., & SCHNEIRLA, T. C. (1942). Mechanisms in conditioning. *Psychological Review*, **49**, 117-134.
- MORGAN, C. L. (1894). *An introduction to comparative psychology*. London: Walter Scott.
- MOWRER, O. H. (1956). Two-factor learning theory reconsidered, with special reference to the concept of secondary reinforcement and the concept of habit. *Psychological Review*, **63**, 114-128.
- RESCORLA, R. A. (1990). Evidence for an association between the discriminative stimulus and the response-outcome association in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 326-334.

- SCHUTZ, S. L., & BITTERMAN, M. E. (1969). Spaced-trials partial reinforcement and resistance to extinction in the goldfish. *Journal of Comparative & Physiological Psychology*, **68**, 126-128.
- SPENCE, K. W. (1956). *Behavior theory and conditioning*. New Haven, CT: Yale University Press.
- THORNDIKE, E. L. (1911). *Animal intelligence: Experimental studies*. New York: Macmillan.
- TOLMAN, E. C. (1932). *Purposive behavior in animals and men*. Berkeley: University of California Press.
- TYLER, D. W., WORTZ, E. C., & BITTERMAN, M. E. (1953). The effect of random and alternating partial reinforcement on resistance to extinction in the rat. *American Journal of Psychology*, **66**, 57-65.
- WOODARD, W. T., & BITTERMAN, M. E. (1976). Asymptotic reversal learning in pigeons: Mechanisms for reducing inhibition. *Journal of Experimental Psychology: Animal Behavior Processes*, **2**, 57-66.

(Manuscript received January 28, 1994;
revision accepted for publication March 13, 1994.)