The Behavior Analyst 1995, 18, 173–185 No. 1 (Spring)

Encore

"Reinforcement" in Behavior Theory

William N. Schoenfeld
Queens College, City University of New York,
and Cornell University Medical College

In its Pavlovian context, "reinforcement" was actually a descriptive term for the functional relation between an unconditional and a conditional stimulus. When it was adopted into operant conditioning, "reinforcement" became the central concept and the key operation, but with new qualifications, new referents, and new expectations. Some behavior theorists believed that "reinforcers" comprise a special and limited class of stimuli or events, and they speculated about what the essential "nature of reinforcement" might be. It is now known that any stimulus can serve a reinforcing function, with due recognition of such parameters as subject species characteristics, stimulus intensity, sensory modality, and schedule of application. This paper comments on these developments from the standpoint of reflex behavior theory.

From its modest beginnings in behavior science, the term "reinforcement" has come to play a central role in modern behavior theory. Like so many others in psychology, the verb "to reinforce," and its cognate nouns and adjectives, were an importation from common usage in which they had seen broad service: reinforced concrete; reinforcing a conclusion; reinforcing a fence; and so on. In acquiring their new dignity in science, they have also acquired some status adjectives like "positive" and "negative." "Reinforcement" itself has won the final encomium of designating an entire branch of behavior theory as "reinforcement theory." All this despite the fact that the term and all its kin lack the clear meanings that theory would desire to lean upon. That fact, though true from the beginning, was indifferently (or so it seems in retrospect) put aside in the enthusiasm of workers over the successes in the practical control of behavior which they were achieving and attributing to "reinforcement." But that fact has now become a rising *obbligato* to all theoretical discussion and can no longer be so comfortably disregarded.

When it entered behavior science, "reinforcement" was intended, in keeping with its common meaning, to convey "strengthening." Pavlov used it that way, as did Skinner and Hull and their followers early and late (e.g., Keller and Schoenfeld, 1950, who first popularized it among a broader scientific public). Of course, all learning theories involve some type of behavior strengthening in various guises and in either explicit or unspoken conventions of reference.[1] So "reinforcement" as it

Reprinted with permission of Transaction Publishers and W. N. Schoenfeld. Schoenfeld, W. N. (1978). "Reinforcement" in Behavior Theory. *Pavlovian Journal of Biological Science,* 135–144. For reprints of this paper and others published in either *Conditional Reflex* or *The Pavlovian Journal of Biological Science,* write to Reprint Department, Transaction, Rutgers—The State University of New Jersey, New Brunswick, New Jersey 08903.

[1] The case is similar to that of the so-called "associationist" theories of learning. Some theorist—I cannot recall who—once wrote that there never has been, and never could be, a learning theory that did not take its origin from an association of some kind. Yet we differentiate among learning theories by their other nuances and conventions, calling some "associationist" and others not. Sometimes we equate "associationistic" with "behavioristic" and "S-R theory"; sometimes we separate Pavlov from Tolman, although both were theorists of the S-S group, and both "associationist." It seems at

developed in the hands of "reinforcement theorists" had to be, and still has to be, differentiated from the strengthenings that are eyed by other theorists, whether they be "cognitive" or "Gestalt" or "mentalist" or "humanist" or whatever. Skinner declared that terms like "reward" and "punishment" could not be of good standing in behavior theory because of their mentalistic connotations, and he offered "reinforcement" as their replacement. It was, in his view, a *technical* word. But in the hands of his followers, "reinforcement" has been returned to its old colloquial self: thus, it has been used in connection with alleged "responses" like "writing a novel," or "getting married," where it can mean only "encouragement" or "reward" in their plain senses (Schoenfeld, 1976).

Having made this full circle under the sponsorship of its most ardent promoter and his followers, it seems timely to re-examine "reinforcement" as a term and as a conception in behavior theory. This *demimonde* may never have been properly rehabilitated for a promenade on the boulevards of behavior theory. If that is so, theory may yet drop it and let it peacefully slip back, as it has tended to do all along, to its simple status *quo ante* in daily nontechnical conversation. As will be seen, little or nothing that is new turns up in the re-examination. But at least some of the implications of what is already known about "reinforcement" will have been assembled.

Pavlov's use of "reinforcement," and his concept of it, was linked to the pairing of a UCS and the signalizing CS. In the pairing, UCS was thought of as "reinforcing" the CS, not the whole reflex CS → R (Gantt, personal communication). In this sense, the UCS "reinforcement" on any single trial of the conditioning procedure may be said to precede the measured response, in contrast to the operant procedure in which, it is customarily said, the "reinforcer" follows R (the cogency of this distinction is questionable, but it is mentioned here only to emphasize that Pavlov's idea of "reinforcement" centered on the stimulus operation in his conditioning procedure, rather than on the measured response; *vide* Schoenfeld, 1972, 1976).

To speak of "reinforcing" CS does not, of course, mean that CS can literally be reinforced. The conditional stimulus is a stimulus, and it is not changed in its physical properties by being linked to a UCS "reinforcer." What is meant is that the "power" of CS is augmented; that is, the "power" to elicit the response under observation, that response which began in the experiment by being the "UCR" to the UCS, more that it was, if ever it was at all, a response to CS. The rise in this "power" of CS is evidenced by an increasing likelihood that the response will be made to it at some magnitude (we may disregard here the usual assumption that the response to CS is the same as the UCR, or some variant of UCR which is on some theoretical grounds acceptable as a category-mate of the UCR; *vide* Schoenfeld, 1966, 1972). Thus, "reinforcement" was both an operation (pairing) and a stimulus (UCS), both a verb (the act of "reinforcement") and a noun (the "reinforcer")—but, then, Western languages easily and often make verbs into nouns, and nouns into verbs.

In any event, "reinforcement" was used by Pavlov, but he gave neither the noun nor the verb the aura and the mystique that they later acquired among operant conditioners. They never dominated his theoretical thinking the way they later did the thinking of operant researchers and theorists.

When Skinner elaborated Thorndike's law-of-effect learning principle into "operant conditioning," the concept of "reinforcement" was placed at the center of the system (Skinner,

---

times as if the nuances of meaning assigned to a given term, and shadings of inference, and conventions of assumption, and such things, more immediately determine the niche a theorist is fitted into than do his actual vocabulary and propositions.

1938, 1969). "Reinforcement" was accorded the status of a "function of the stimulus," and was made the *vis viva* of operant learning in all its applications and variations: response acquisition and "shaping," behavior therapy, behavior modification, and all the rest. It was said that "rewards" and "punishments" were vulgar notions of how operant behavior is controlled; their replacement by "reinforcement" was urged on the grounds of technical, operational, and analytic clarity, and therefore of greater advantage to the progress of behavior science and behavior theory.

The core idea of Skinner's new system was that "reinforcement" could be applied to responses. If this had meant only that the general idea of "strengthening" was applicable to responses, then taking over the term "reinforcement" from Pavlov's vocabulary might have been reasonable. But in its new address as a technical term it was implied that there was a *technical* identity of process or operation or function or whatever in the two cases, and that was, at the least, a dubious equation whether on the theoretical or the experimental level. Even Skinner did not believe in the identity. This is seen in the way he, as compared with Pavlov, treated the "reinforcement" of stimuli. The emerging feature of that treatment was the "discriminative stimulus," the "S^D." He resisted the description of discrimination training, the building of "stimulus control" over operant responding, as a procedure for "reinforcing" the S^D. Such a description would have located the matter within Pavlov's domain, however sketchy that was. Instead, Skinner spoke of the S^D as only "setting the occasion" on which the response would be reinforced, or would be eligible for reinforcement. If the S^D could be said in any way to be reinforced, it was only through the mediation or intervention of the response. That Skinner did not wish to identify wholly with Pavlov, and thereby put operant conditioning into some niche within classical reflex theory, is apparent in his debates with Konorski and Miller (1937a, 1937b; Skinner, 1937). The refusal to regard [$S^D \cdot R$] as a case equivalent to [$CS \rightarrow R$], and the rejection of every effort to treat S^D and CS in the same categorical terms, was the outcome (and the evidence) of Skinner's new use of "reinforcement." That use was not congruent with Pavlov's. At the same time, he affirmed his allegiance to reflex theory, and signified it by writing the "operant" as a reflex [$s \cdot R$] even though the stimulus was only a lowercase s and could (he held) be neglected once control over R was assumed by the "reinforcement" (Skinner, 1931, 1938). By his refusal, and by his ambivalent embrace of reflex theory, Skinner planted a seed of conflict at the very beginning of his work. That seed has flowered openly in recent years. The mainstream of operant work and theory is no longer in the reflex tradition, although it still sees itself as part of a "behavioristic" tradition.

Operant "reinforcement" was soon divided into two subcategories: "positive" and "negative," or "appetitive" and "aversive." Although opposed in name, they were not ultimately regarded as opposed in action. Briefly, the subcategories were defined as they were in the sense of the original Thorndikian law-of-effect: a "reinforcer" was positive/appetitive if a response that produced it was more likely to be repeated, and negative/aversive if a response that produced it was less likely to be repeated (or a response that terminated it was more likely to be repeated). Thorndike used "satisfaction-discomfort," but their replacement by the allegedly technical term "reinforcement" had changed (and was to change) little, if anything, in the meaning. Even Thorndike's well-known decision that while an outcome of "satisfaction" could strengthen a response (that is, a response of the type later called "operant"), the outcome of "discomfort" would not literally subtract strength from a learned response (though it might suppress it temporarily), was echoed by later operant findings (Skinner, 1938; Estes, 1944; Keller and Schoenfeld, 1950).[2] Not only were both laboratory practice and theoretical discussion left unchanged by the substitution of "reinforcement" for Thorndike's terms, but over the years the writing of operant conditioners has lost Skinner's discipline and has taken on a looser Thorndikian style.

It might be noted almost as an aside that difficulties arose early about how the adjectives "positive" and "negative" were to be used. To some writers (e.g., Keller and Schoenfeld, 1950), "positive reinforcement" meant that a "positive reinforcer" was being applied to the organism, and "negative reinforcement" that a "negative reinforcer" was being applied; but the application could be either that of supplying or of withdrawing each type of reinforcer when the stipulated response was made. To other writers, "positive" and "negative" did not mean the type of stimulus being applied to the response, but rather the effect upon the response itself: thus, "positive reinforcement" meant any application which increased response "strength" regardless of the stimulus operation being performed and so on. Still other usages were favored in various quarters, depending on which aspects of the experimental operations or the subject's response or the response's consequences were selected for characterization as "positive" or "negative." This has never been more than a matter of whimsy, however, and does not intrude upon really serious considerations regarding the term and concept of "reinforcement."

When Pavlov began work on his "conditional reflex," certain conventions and assumptions were interlaced with his experimental plans and procedures. The same is true, of course, of every mind in every time, both in and out of science. Our particular concern with Pavlov's case is that many, if not all, of those conventions and assumptions still figure in the thinking of contemporary behavior scientists when they should, on the evidence, have been discarded or modified by now. Pavlov himself discarded some that he started with because his insight grew as his work progressed, but even of those which he discarded several have persisted among theorists to this day.

One idea Pavlov had at the start was that a "reinforcement" could be "unconditional." That apparently carried along several other notions: for example, that a given "reinforcement" could be "natural" to a particular organism like the dog and be an unlearned part of a biological reflex; this, in contrast with a "neutral" stimulus which (though it might be the "unconditional" stimulus of a reflex with its own "natural" response) did not originally elicit the same response as the "unconditional" stimulus. Skinner's (1938) statement of the "Law of Conditioning of Type S" emphasized this: "The approximately simultaneous presentation of two stimuli, one of which (the 'reinforcing' stimulus) belongs to a reflex existing at the moment at some strength, may produce an increase in the strength of a third reflex composed of the response of the reinforcing reflex and the other stimulus" (p. 18). Some contemporary authors of texts still hold to the idea that something can be "inherently reinforcing" (e.g., Reese, 1978, pp. 18, 55). Pavlov himself had not gone far into his behavioral research before he asked just what aspects of the "unconditional" stimulus were really unlearned or natural. An "unconditional" stimulus such as food is actually an "object" having many stimuli emanating from it, each of which (say, the look and feel and smell and taste) had to be learned by his dog. It is for this reason that an experimen-

[2] In the recent rise of "aversive control" as a major research area for operant conditioners, the recognition has been growing that the simplicity once sought regarding the relative effects of "positive and negative reinforcers" is not attainable. This is probably more in keeping with the "old-fashioned" views. A closer analysis of the real action of all "reinforcers" must surely follow.

tal animal newly arrived in a laboratory needs to be adapted to the regimen there, including the food; for the same reason, people accept accustomed foods and bypass or reject others. Without training, neither we nor the laboratory dog would even recognize a proffered substance as "food," and make no move to eat it even when hungry. Pavlov saw this and emerged willing to take the animal as he found him, although there were still many unspoken conventions to be met before the animal would be acceptable: he had to be healthy, have his senses, be "normal" in behavior and not too frightened by the laboratory, friendly, and so on, since otherwise he would not be *useful* for the studies Pavlov had in mind. Each of these requirements is, obviously, an area of added behavioral and physiological interest, but for Pavlov's purposes they were parameters to be kept relatively constant. So long as the salivary response could be depended upon with the stimuli applied, so long as this base reflex was "in," so long as the behavioral repertoire included it, the work of conditioning the response to a new stimulus could begin. The "unconditional" stimulus was really another "conditional stimulus" (it could as well be called S_2 or any other nonpartisan name), but that did not matter so long as pairing it with the "conditional stimulus" (say, S_1) produced a transfer to the latter of some elicitative power over R.

Among the parameters of the conditioning procedure which seemed only common sense in those early years, and which were manipulated only to allow the conditioning work to proceed, was that of "motivation." It was taken for granted that a dog had to be "hungry" to perform properly, but, as we know, the analysis of "hunger" or any other "motive" or "drive" has since then become a serious concern of behavior science. A "drive-establishing" operation may "prepare" the organismic system for "reinforcement" in the sense that when the "reinforcement" is delivered it will actually be "reinforcing." But every such operation is also acknowledged to affect an organism's internal mechanisms, and its relations with its environment, in many ways. Playing off Skinner's description of the "S^D" as "setting the occasion" on which a response may be reinforced, we might say that a "drive-establishing" operation is one that "sets the system." It does so in many functional ways, including sensitivity to the proffered "reinforcement."

"Secondary reinforcement" has its historical root in Pavlov's "higher-order conditioning," but the same theme was already implicit in Pavlov's hesitation over the "unconditional" status of a "reinforcement" like food. What emerged from the higher-order conditioning procedure was the observation that stimuli which were unquestionably "conditional" could serve as "reinforcements." Many later workers in the operant camp were unprepared to take the theoretical line (did it seem to them overly bold?) that all "reinforcers" fall alike, and preferred to maintain the dual classification of "primary reinforcements" and "secondary reinforcements" (e.g., Wike, 1966). Synonyms gradually accumulated for the two classes, depending upon the examples cited by authors, and upon the authors' personal histories: unconditional or primary or vital or biological or unlearned or innate or natural *versus* conditional or secondary or nonvital or social or learned or derived. Textbook writers struggled valiantly to differentiate the two classes, but teachers in the elementary classroom never could satisfy the brighter of their puzzled students. Some theorists (e.g., Miller, 1951) tied their concept of "reinforcement" to that of "drive," holding, say, that a "drive-establishing" operation creates a potentiality for "reinforcement" by setting up a "need"; that the "reinforcement" acts to reduce this "need" through its undoing or reversing what the "drive" operation established; and, that it is the reduction of "need" which makes the reinforcer "reinforcing." For these theorists, the existence of "secondary reinforcements" meant that drives had to exist from which such "reinforcements" could draw their functional power. If the drives could not be demonstrated to exist, they had to be postulated; and, to conform with the learned character of "secondary reinforcements," their corresponding "drives" had also to be secondary or learned or acquired or derived. This way of handling "secondary reinforcement" dealt motivation theory a wound from which it has not yet recovered. It demanded a trafficking in infinites: since the number of stimuli which could be made into "secondary reinforcements" is indefinitely large, and since each demanded a "drive" to account for its function, the number of "learned drives" which had to be postulated was also indefinitely large. Perhaps it is that which Skinner wished to satirize when he proposed that a king, calling "A horse, a horse, my kingdom for a horse!," had a "horse-getting drive."

Theorists who lean toward a dual classification of "reinforcements" into primary and secondary try to find a reliable criterion to distinguish them. When the subject arrives in the laboratory, he is already under the potential control of one or more "reinforcements," and these therefore appear to be a "natural" part of his behavioral schema, and on that account "primary." Moreover, the laboratory worker has his conventional notions about how animals are with their "needs." Apart from these facts, the distinctions that have been advanced to segregate "primary" from "secondary" reinforcements can be surmised from the synonyms mentioned earlier which have accumulated for them. None of these distinctions, however, endures well under scrutiny. The most durable of these criteria, and the one upon which most theorists have converged, is that of extinguishability: a "primary reinforcement" is supposed not to be extinguishable with continued use (though it may temporarily satiate), whereas a "secondary" one would extinguish with continued use unless it is backed up by at least occasional reassociation with a "primary," whereby its power is renewed. In the end, this criterion fares no better than the others. But, unhappily, it has led its proponents to a futile experimental search for those stimulus aspects of a "reinforcement" which were doing the actual reinforcing, and for the actual site of reinforcement, that is, just *where* the reinforcement was going on. For example, if food were shown to a subject, but he was not permitted to take it into his mouth; or, if he were allowed to chew it, but not to swallow it; or, to swallow, but not have it reach his stomach; or, to reach the stomach, but not to be digested; and so on down the track—what parts of the stimulational chain would cease to "reinforce" with repeated trials, and where would the reinforcing effect go on and on even if the later parts of the ingestive-digestive sequence were not reached?

The upshot of all this thought and effort regarding possible distinctions between "primary" and "secondary" reinforcement was a disillusion with the problem. Where once the literature was crowded with studies and discussions of the topic, there is now a disappointed and exhausted silence. It seems now safe to say that the two cannot be distinguished. When an experimental animal is chosen, he is taken as he is. A stimulus is a stimulus, and a "reinforcer" is as "reinforcer" does. This conclusion is related to another issue which has been the focus of some discussion in the literature, namely, the nature of "reinforcement"; that issue is returned to later.

Two matters, perhaps a bit digressive, might be noted at this point. The first has to do with the parameters of "secondary" reinforcements; the second, with the actual procedure for imparting "secondary" reinforcing power to a hitherto "neutral" stimulus or object or event. Both matters involved testing or measuring the power of a "secondary" reinforcer after it was established, and three methods for doing this were in general use among operant conditioners. One measure was the power of a "secondary" reinforcer to condition a new response (akin to the operant conditioning power of a "primary" reinforcer; or akin, in Pavlov's case, to the power to "reinforce" CS); the second measure was the power to retard extinction (for which Bugelski's experiment of 1938 served as a model); and, the third, the power to reinstate responding when introduced into the course of extinction (which Skinner had demonstrated, 1938). These might in theory be variants of a single measure, but they were individually used by researchers into the parameters and procedures of "secondary" reinforcement.[3]

For "primary" reinforcement, such parameters as number, weight, amount, size, type, sense modality and delay, were studied (e.g., Hull, 1943; Zeaman, 1949; Zeaman and House, 1950), but today the meaning, and sometimes even the reality, of such variables is not secure. This is the case with the so-called "delay of reinforcement" variable which appears to be a reasonable parameter of response reinforcement procedures until the fact of continuity in the behavior stream is acknowledged and converts "delay" into a problem of response chaining. For "secondary" reinforcement, the popular parameters of study were number of pairings with a "primary," and time separation from the "primary." Thus, Bersh (1950) showed that this number-of-pairings function resembled the accepted number-of-reinforcements function for original conditioning, and that this time-separation function was in close accord with accepted CS-UCS interval functions. While this was true for "positive" reinforcement, the corresponding number-of-pairings and time-separation functions for "negative secondary reinforcement" were significantly different: the number-of-pairings function went through a maximum, and the maximally effective separation time was perhaps a whole order of magnitude greater than in the "positive" case (Libby, 1951), The latter findings have never received the attention they seem to deserve (e.g., Hall, 1976) from theorists interested in what is today called "aversive control." Again, as in the "positive" case, such variables in the "negative" case lost some of their import over the years. The omnipresent "reinforcement schedule" was early recognized as a major parameter of all such behavioral studies (Skinner, 1938; Ferster and Skinner, 1957), but even this matter has come in for reconsideration and recodification by contemporary theory (Schoenfeld and Cole, 1972, 1975). In any case, the interests of researchers have taken new directions which have little relation to the original problems of the properties of "primary" and "secondary" reinforcement, and of possible distinctions between the categories.

The problem of establishing a "secondary" reinforcement has taken a somewhat similar historical course. When the period of active operant research opened, it had not occurred to anyone, or no one had clearly recognized, that any difficulty lay in store. It was thought that the only requirement for a stimulus to become a "secondary reinforcer" (or, for that matter, a "higher-order" CS) was, as Hull (1943) put it, a close and consistent relation to a "primary." Things came into better focus when a certain experimental procedure which did incorporate such a relation failed to produce a "secondary reinforcer" (Schoenfeld, Antonitis, and Bersh, 1950). Upon rethinking their procedure, and after reviewing the available literature on es-

[3] During this same period, some of these workers were also tackling the separate but related question of the generality of a "secondary" reinforcer once it was established. That is, whether, after establishment, such a reinforcer could be used on a new response, under a new "drive," or with both response and drive shifted. The findings indicated that complete generality could be expected (that is, capacity to condition a new response under a new "drive"), but interest in this question has also faded since those years.

tablishing "secondary reinforcers," these workers suggested that the requisite training procedure was the same as that for producing a "discriminative stimulus," or S^D. This suggestion caused a flurry of controversy, and became known in the literature as the "discrimination hypothesis of secondary reinforcement." Several investigators attempted to show that the suggestion was incorrect, but their attempts never really succeeded, while opinions about the issue have continued opposed. Two things, however, may still be said of the "hypothesis." First, it has remained since its time the only explicit suggestion of what was required for establishing a "secondary reinforcer." If that suggestion is wrong, no other operational statement has replaced it either to guide an experimenter in a practical way who may want to create such a "reinforcer," or to guide theorists in considering how such a "reinforcer" relates to any other stimulus. They are left with only the open-ended "close and consistent" specification, and the problem is returned perhaps to the simplistic level of accepting a "reinforcement" whenever, wherever, and however one can find it. Second, as with the parameters of "secondary reinforcement," the procedure for establishing one in the first instance seems to have lost its interest for contemporary theorists. It has been bypassed in their march to other issues that are currently more popular. Why that has happened here, and indeed how it happens (as it often does) with any issue in any science at any historical stage, is not obvious. Trends of popularity in science are so far a concern only of historians of science. One day they will also be of concern to a behavior scientist who undertakes to study, as he would the behavior of any organism, that of the scientist who, against his own complex background and his social group's, is responding in his immediate surroundings to an accumulating mass of new empirical information.

It is a commonplace in lay experience that several sorts of "reinforcement" ("rewards") are, or can be made, effective with an animal like the dog. Thus, when "reinforcement" became an arena of thought for behavior scientists, a seemingly natural question emerged: what is the "nature of reinforcement?" What do all "reinforcers" have in common which enables them to act as such? Answers were not long in being proposed, perhaps as many as a dozen of them (cf. Glaser, 1971). Some were fanciful and could never be translated into testable terms; some were paraphrases of conventional beliefs and myths of the culture, which seemed, but only briefly, to be testable in the laboratory; and, some were purely pragmatic, rather than rational. A few of the more well-known of these proposals might be mentioned here, even if only cursorily.

One of the pragmatic views was that of Skinner, for whom a "reinforcer" was as reinforcer does. In working with any organism, human or animal, a "reinforcer" is sought out which works for that subject and which is convenient to use; the scientist then proceeds without further ado. On a simple level, this view has the virtue of being incontrovertible. Moreover, it gave the lie to the question of the "nature" of reinforcers by merely ignoring it. If the question is a reasonable one, however, it may be handled in a way that rationalizes Skinner's simple empiricism. That way was offered early by Guthrie (1930, 1935, 1952) within the framework of reflex theory as he saw it. He gave it various statements because, as he said, he was trying to reach different audiences however it was necessary to make his point; but none of his efforts, nor all together, in the end convinced his colleagues (Mueller and Schoenfeld, 1954; Hilgard, 1956; Hilgard and Bower, 1966). Guthrie was applying to the problem of "reinforcement" the Leibnizian principle of sufficient reason: since, in a reflex, S is the "cause" of R, whenever R occurs we must suppose that S has occurred; if the rate of occurrence of R is changed, we must suppose that the rate of occurrence of S has changed; if, in a given experimental arrangement or procedure, a stimulus change (the "reinforcer") following an R changes the rate of R, we must say that the "reinforcer" is making the S of the reflex recur more often. Guthrie voiced this reasoning in various terms, such as the "postremity principle," or "protecting" the stimulus of the reflex, and so on. He intended only to say that if a given S once initiates an R reflexly, then the organism will make the same R if put back into the same initiating circumstance. Similar causes produce similar effects; similar effects come from similar causes. Science can say no less, and Guthrie was surely correct from the standpoint of reflex theory. The difficulty in practical terms was to say precisely "how much" of the original S had to recur for R to recur; how much a "reinforcer" had to change the organism's stimulus ambience in order to "protect" the original initiating circumstance, that is, in order to keep it unchanged so that, if the organism were put back into it, the same R as before would occur. Confronted with the demand by his fellow theorists for a practical statement, Guthrie admitted that, for the present at least, his answer would have to be an empirical one: a "reinforcer" would have to be tried, and if it worked, so be it; if it failed, we would know it was insufficient. This position was, of course, the same as Skinner's (though with a rationale provided), but the identity was overlooked by his colleagues, including his critics, while supposed differences between these two theorists were being debated.

Another, and perhaps the most popular, proposal of a "nature" that all reinforcements share is that they are "need reducing." The idea is that conditioning or learning requires that the subject be in a state of motivation, and that such a state involves a "need." Drive establishment is translated into "need establishment," and, as said earlier, a "reinforcement" is seen as acting to undo, or to reverse, the "need" that has been established. Hull (1943), who was among the theorists taking this position, even argued that Pavlovian "reinforcement" shared this property with the operant or instrumental case, though it took much convoluted (and in the end unconvincing) argument and diagramming by him to force the UCS into that mold. The idea of "reinforcement" as "need reducing" has been most popular among the more "behavioristic" theorists, probably because it seems to eschew any subjective or mentalistic attribute of "reinforcement." "Need" has a biological ring to it, a hint of notions like "tissue needs for healthy functioning"; linked to that is the aforementioned belief that the ultimate locus of "primary reinforcement" can be tracked down somewhere in the soma. But that locus proved to be only a will-o'-the-wisp, and the concept of "need" in a physical-chemical machine (which is how behavioristic theoreticians commonly view an organism) seems misplaced. Machines do not "need" anything. When, for example, we say in daily speech that our automobile "needs" oil, we do not mean that *it* needs oil: it will work in accord with certain physical principles with oil, and in accord with others without oil. The car itself "needs" nothing. It is *we* who need the car to have oil, since we wish it to function in one way and not another. So with a human or animal subject deprived of food, say: as a physical-chemical system, it will work one way with food, another without; it "needs" nothing, but we do who wish it to live and to function for *our* sakes. To a true behaviorist, "need" is only a socially inspired and possibly poetic response from laymen peering over his shoulder as he works, and directing their response to his experimental subject when they should be directing it at him instead.

Akin to the concept of "need," is that of "homeostasis" as the definitive element of the "nature of reinforcement." The model is the chemist's buffered solution, and examples are commonly cited from the mammalian body in which homeostatic functions of organs and blood are known. From these, the model is extrapolated to behavior: thus when a "drive" is established, say by food deprivation, the organism's "homeostatic" balance is said to be upset, and his behavior is viewed as being directed toward restoring that balance; the food is said to reverse the upset, to restore the balance *quo ante,* and it is that which makes of food the "reinforcer" that it is. Some theorists have extended the "homeostatic" concept to behavior universally, even to speaking of a tic as the patient's effort to restore himself to his prepathological status, or to establish a new normative balance. The difficulty with defining "reinforcement" by reference to homeostasis does not lie in the phenomenon of homeostasis itself, the chemistry of buffers, or in hematology. The difficulty is rather with the extrapolation of the homeostatic model to any specific behavior, such as eating. Some years ago, a cogent critique of this extrapolation appeared in the literature. I cannot recall the author, but among his points was the fact that it is necessary to know, or to have some way to assess, the prior departed-from baseline behavior, or baseline of quiescence, to which the homeostatic function is attempting to restore the organism. Otherwise, a "homeostatic" explanation of any specific instance of observed reinforcement and responding can only be *ad hoc,* without theoretical power, and unable to guide us in the practical and predictive control of behavior. His criticism seems still to be correct, and equally applicable to related concepts like that of "drive" as a raised "tension" or "arousal" or "activation," with the virtue of a "reinforcement" being that it is "tension reducing" and acts to restore the predrive "relaxation"; or the concept that "drives" which are established by "aversive stimulation" are destabilizing, and are accompanied by the state or the "emotion" of "anxiety," and that the "reinforcement" effect is accomplished by "anxiety reduction."

A similar critique may be framed for the suggestion that the "nature of reinforcement" involves the making of a "consummatory response." Thus, it is said, it is the act of eating food that is "reinforcing," and not the food itself. Proponents of this view of the "nature of reinforcement" have carried out experiments to show that the *size* of the food pellet is not critical, nor any physical aspect of the food, but rather that the offering be actually eaten. Eating is the "consummatory response" which confers "reinforcement" value to the food. The difficulty with this suggestion is to identify the actual response which is the "consummatory" one: is it seeing the pellet, approaching it, picking it up, putting it into the mouth, chewing, swallowing, stomachic contraction upon it, or whatever down the line? In an organism's continuous behavior stream, every response or none can be said to be "consummatory." There is a nuance of "finality" to the term "consummatory," but appeal to that nuance can only be *ad hoc,* and after the fact. What is observed to be "reinforcing" can be claimed to be "consummatory," but in advance of an observed "strengthening" of a response no such claim is dependable.

Still another effort to specify a special "nature" that runs through all "reinforcers" is the so-called Premack Principle which looks to responding, rather than stimuli, as the vehicle of "reinforcement." In the operant case, the principle states, a preferred response can serve as "reinforcer" for a less preferred, whereas the latter cannot "reinforce" the former. In practice, the indicated procedure under the principle is to make the subject's access to a preferred response contingent upon his making the required, but less preferred, response; that is, to give the subject an opportunity to do something he desires to do, if he makes a response that the experimenter desires and stipulates as required. The difficulty with the principle is to evaluate the preference for one response over another, since without such a relative evaluation the principle has no meaning or use. This requires that the "responses" being employed, and their boundaries, be defined, so that it is possible to detect and to count or otherwise measure when *a* response of one sort or the other has occurred and thereby be able to compare "preferences" by some criterion of comparability (which also needs to be specified). Is relative preference to be estimated by the proportion of times spent with each response in a free-choice situation—if so, does a minute at one response equal, in some sense, a minute at the other, or are these equalities or inequalities judgeable only after the fact? Since a "response" like bar-pressing is a punctate one of as short a duration as we wish, how is it to be compared with a "response" like activity wheel turning which is durative—the former has a rate, but does the latter, and the latter has duration, but does the former? If relative frequencies of two responses are to be the basis for evaluating relative preference, can the boundaries of a response be determined, or cannot any "response" or segment of behavior be dissected into smaller and smaller components, or incorporated as components of larger and larger units of behavior, with seeming increases and decreases in the frequency of each response resulting from these arbitrary assignments? A number of workers have claimed success in applying the principle to obtain control over behavior in practical situations like psychotherapy and classroom teaching. But, while their successes may be genuine, it is not the Premack Principle which can be credited with them.

Other suggestions, less popular than those mentioned, about the "nature of reinforcement" have been made. One of them, however, contains the resolution of the problem. It has been around for a long time, passing under different guises in different connections, and perhaps on that account has been slighted except by some few theorists like Guthrie who voiced it in his particular terms. Simply put, it is that *any* stimulus can act as a "reinforcer": there is no special class or group or characteristic of "reinforcement" which sets it apart from other stimuli, but rather that all stimuli can act so depending upon their intensity, static and dynamic patterning, locus of application on the organism, sense modality, and still other parameters. In the early literature, this fact underlay even such so-called "drives" as "curiosity" and "novelty" and "exploration," the considerable strengths of which were noted when they were pitted against such standbys as hunger and thirst and sex (e.g., Warden, 1931). The tests involved stimulus consequences for a subject's behavior which were changes from its present ambience; they were changes which supposedly pricked "curiosity" and "exploration," but stimulus changes they were nevertheless in their direct operational significance. In more recent years, in the "operant" or "instrumental" context, the fact has been demonstrated many times that any stimulus change can act as a "reinforcer" for a response upon which it is made contingent: allow a designated response to turn on a light or a tone, to raise a curtain so a fellow organism can be observed, to bring the responder into a novel environment, and so on and on—any of these consequences can be shown to be capable of "reinforcing" the response producing it (*vide,* Hall, 1976, esp. pp. 237–239).

If it is correct to conclude that any stimulus change can be "reinforcing," we can leave behind the seemingly futile search for different *classes* of stimuli called "reinforcing" and "nonreinforcing." There would still remain the empirical problem noted earlier in connection with Guthrie's and Skinner's views of "reinforcement": that is, to pin down the precise characteristics of the stimulus changes which secure for us those behavioral effects that we attribute to "reinforcement."

We might summarize now the burden of these remarks regarding "reinforcement."

Pavlov used the term descriptively in his version of behavioral reflex theory, intending by it the action of a stimulus from an established reflex in imparting to a second stimulus the capacity to substitute for the first as an elicitor of the response. When it was taken over into operant behavior theory, which originally saw itself as falling within the compass of reflex theory, the term was intended to be a technical one to replace such allegedly mentalistic terms as "reward" and the "satisfaction" which figured in Thorndike's Law of Effect. Of original interest in operant conditioning was a "contingency" relation between a response and its "reinforcer." More lately, the interest of operant conditioners has broadened to include "noncontingency," but "reinforcement" has still retained for them its key role in all functions of response acquisition and extinction (Schoenfeld et al., 1973). In retrospect, "reinforcement" never did become either a technical term or an analytic one. In practice, its actual referent became only complex, including an indefinite number of experimental operations, and the behavioral observations preceding and following those operations which comprise the data of "conditioning." Moreover, current operant behavior theory is no longer part of general reflex theory, and this historical development is reflected in the meanings and uses of such terms as "reinforcement" and "response."

It was early, and is still in many quarters, believed that only certain stimulus changes ("behavioral consequences") could produce operant learning, that is, could serve as "reinforcements." Theoretical conjectures and disputes arose concerning the "nature of reinforcement," that is, whether all "reinforcers" shared some property by virtue of which they secured their behavioral effects. Of all the suggestions made, none survives critical examination save one: that any stimulus change can be a "reinforcer" if the characteristics of the change, and the temporal relation of the change to the response under observation, are properly selected. This conclusion is supported both by a rational argument within the framework of reflex theory, and by experimental evidence from a variety of operant conditioning situations. As yet, this conclusion shares one serious difficulty with its rejected alternatives: to wit, no way is so far known to derive rationally, and in advance of actual test, the characteristics of stimulus change which will produce "reinforcement." Accordingly, the actual uses of "reinforcement" are still in the pragmatic realm where an experimenter makes his selections and applications on the basis of common experience. His success is judged after the fact by the demonstration that his subject has "learned" or "been conditioned." But if the investigator takes his stand upon reflex theory, he will, first, specify his experimental operations in the terms of physical science; and second, he will treat "conditioning" in stimulus-response terms. In the latter treatment, he will rely upon the Leibnizian principle of sufficient reason, and upon the same principle of cause-and-effect that is indispensable in every natural science.

REFERENCES

Bersh, P. J. (1951). The influence of two variables upon the establishment of a secondary reinforcer for operant responses. *J. Exp. Psychol. 41,* 62–73.

Bugelski, R. R. (1938). Extinction with and without sub-goal reinforcement. *J. Comp. Psychol. 26,* 121–134.

Estes, W. K. (1944). An experimental study of punishment. *Psychol. Monogr. 57,* No. 263.

Ferster, C. B. and Skinner, B. F. (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Gantt, W. H. Personal communication.

Glaser, R. (Ed.). (1971). *The nature of reinforcement.* New York: Academic Press.

Guthrie, E. R. (1930). Conditioning as a principle of learning. *Psychol. Rev. 37,* 412–428.

Guthrie, E. R. (1952). *The psychology of learning.* New York: Harper.

Hall, J. F. (1976). *Classical conditioning and instrumental learning.* Philadelphia: Lippincott.

Hilgard, E. R. (1956). *Theories of Learning* (2nd ed). New York: Appleton-Century-Crofts.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of Psychology.* New York, Appleton-Century-Crofts.

Konorski, J., & Miller, S. (1937a). On two types of conditioned reflex. *J. Gen. Psychol. 16,* 264–272.

Konorski, J., & Miller, S. (1937b). Further remarks on two types of conditioned reflex. *J. Gen. Psychol. 17,* 405–407.

Libby, A. (1951). Two variables in the acquisition of depressant properties by a stimulus. *J. Exp. Psychol. 42,* 100–107.

Miller, N. E. (1951). Learnable drives and rewards. In Stevens, S. S. (Ed.), *Handbook of experimental psychology.* New York, Wiley.

Mueller, C. G., Jr., & Schoenfeld, W. N. Edwin R. Guthrie. (1954). In Estes W. K. et al. *Modern learning theory.* New York, Appleton-Century-Crofts.

Reese, E. P. (1978). *Human behavior: Analysis and application* (2nd. ed.). Dubuque: Wm. C. Brown.

Schoenfeld, W. N. (1966). Some old work for modern behavior theory. *Cond. Reflex, 1,* 219–223.

Schoenfeld, W. N. (1972). Problems of modern behavior theory. *Cond. Reflex 7,* 33–65.

Schoenfeld, W. N. (1976). The "response" in behavior theory. *Pavlov. J. Biol. Sci. 11,* 129–149.

Schoenfeld, W. N., Antonitis, J. J., & Bersh, P. J. (1950). A preliminary study of training conditions necessary for secondary reinforcement. *J. Exp. Psychol. 40,* 40–45.

Schoenfeld, W. N., & Cole, B. K. (1972). *Stimulus schedules: The t-t systems.* New York: Harper and Row.

Schoenfeld, W. N., & Cole, B. K. (1975). What is a "schedule of reinforcement"? *Pavlov. J. Biol. Sci. 10,* 52–61.

Schoenfeld, W. N., Cole, B. K., Lang, J., & Mankoff, R. (1973). "Contingency" in behavior theory. In F. J. McGuigan & D. B. Lumsden (Eds.), *Contemporary approaches to conditioning and learning,* V. H. Winston: Washington, D. C.

Skinner, B. F. (1931). The concept of the reflex in the description of behavior. *J. Gen. Psychol. 5,* 427–458. See also the added foreword to this paper when reprinted in B. F. Skinner, *Cumulative Record,* New York: Appleton-Century-Crofts, 1959, pp. 319–320.

Skinner, B. F. (1937). Two types of conditioned reflex: A reply to Konorski and Miller. *J. Gen. Psychol. 16,* 272–279.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1969). *Contingencies of reinforcement.* New York: Appleton-Century-Crofts.

Warden, C. J. et al. (1931). *Animal motivation.* New York: Columbia University Press.

Wike, E. L. (1966). *Secondary reinforcement.* New York: Harper and Row.

Zeaman, D. (1949). Response latency as a function of amount of reinforcement. *J. Exp. Psychol. 39,* 466–483.

Zeaman, D., & House, B. J. (1950). Response latency at zero drive after varying numbers of reinforcements. *J Exp. Psychol. 40,* 570–583.