NEGATIVE REINFORCEMENT IN APPLIED BEHAVIOR ANALYSIS: AN EMERGING TECHNOLOGY

BRIAN A. IWATA
UNIVERSITY OF FLORIDA

Although the effects of negative reinforcement on human behavior have been studied for a number of years, a comprehensive body of applied research does not exist at this time. This article describes three aspects of negative reinforcement as it relates to applied behavior analysis: behavior acquired or maintained through negative reinforcement, the treatment of negatively reinforced behavior, and negative reinforcement as therapy. A consideration of research currently being done in these areas suggests the emergence of an applied technology on negative reinforcement.

DESCRIPTORS: aversive stimulation, avoidance, escape, negative reinforcement

Research published in the Journal of Applied Behavior Analysis (JABA) has, for 20 years now, demonstrated how knowledge about environment-behavior interactions, particularly those involving response-contingent events and correlated stimuli, may be used for the benefit of individuals and the larger society. In doing so, applied research has also made significant contributions to the general science of behavior by providing extension and external validation of experimental findings from the basic research laboratory (Baer, 1978).

Along with the development of the applied field and its expansion into a number of areas in which the outcome of an experiment often has immediate social implications (e.g., business and industry, developmental disabilities, education, medicine, mental health, public affairs), there has been growing concern of a widening gap between basic and applied behavior analysis. Critics (Deitz, 1978; Pierce & Epling, 1980) have indicated that the emphasis of contemporary applied behavior analysis has shifted away from the study of conditions that produce change to the production of change per se, and that "relevance to basic principle," a supposed characteristic of applied behavior analysis (cf. Baer, Wolf, & Risley, 1968), is reflected less and less in the research that journals such as JABA publish. The general accuracy of these criticisms, as well as their basis and implications, will continue to be the subject of periodic debate (Baer, 1981; Cullen, 1981; Michael, 1980). Nevertheless, it is possible at this point in the development of our field to identify specific and well-established areas of basic research for which little parallel exists in the applied literature, and vice versa.

The thesis of this article is that research on negative reinforcement provides one of the clearest and most immediately relevant examples of a case in which consideration, replication, and extension of basic research would benefit the applied area. Along with positive reinforcement and punishment, negative reinforcement has long been considered one of the elementary principles of behavior. A voluminous amount of research on negative reinforcement exists in the basic literature (see reviews by Bolles, 1970; Herrnstein, 1969; Hineline, 1977, 1981, 1984; Hoffman, 1966; Schoenfeld, 1969; Sidman, 1966), and its inclusion as a distinct topic in texts on experimental analysis (e.g., Honig & Staddon, 1977) justifies its status as a major organizing principle. For example, acquisition, maintenance, extinction, and stimulus control all have been studied using negative reinforcement as the operant mechanism of interest.

Sandler and Davidson (1973) reviewed some of this basic research and discussed its relevance to the
development and treatment of pathological human behavior. They concluded that "... the escape and avoidance paradigms are still plagued by a number of unresolved issues ..." (p. 254) that they hoped would be clarified by additional basic research and extension to the world of humans. Since that time, a number of investigations on negative reinforcement with humans have been conducted, yet a systematic and comprehensive body of applied research still does not exist. Consider the two most recent texts on aversive control with humans (Axelrod & Aspe, 1983; Matson & DiLorenzo, 1984). Both provide thorough discussion of topics such as positive reinforcement, extinction, time-out, response cost, and contingent aversive stimulation. Thus, one might expect that these texts would be the most likely sources of information on negative reinforcement as well, but this is not the case. One text (Axelrod & Aspe, 1983) devotes less than a half dozen of over 300 pages to the topic of avoidance, and the discussion always is limited to avoidance as a side effect of punishment. No mention is made of escape or avoidance as directly produced performances. The second (Matson & DiLorenzo, 1984) describes the hypothetical features of escape and avoidance training on two pages but does not cite any applied references.

The relative absence of integrated material on negative reinforcement with humans raises several questions concerning generality and utility. Is human behavior relatively insensitive to contingencies of negative reinforcement? Are naturalistic human situations typically characterized by the absence of stimuli that can function as negative reinforcers, or opportunities to escape from or avoid these stimuli? What types of performances are likely to be acquired through negative reinforcement? Finally, do procedures based on the application of negative reinforcement, unlike those based on positive reinforcement and punishment, have little therapeutic or pragmatic value?

For the past few years my students, colleagues, and I have been conducting a series of investigations in two areas—self-injurious behavior and pediatric feeding disorders. Curiously, these very different problems have brought us into direct contact with situations involving the use of negative reinforcement and have forced us to consider more generally the relevance of negative reinforcement in applied behavior analysis. Our experience and our examination of the basic and applied research suggest that the answer to each of the above questions is "No." In fact, it appears that negative reinforcement plays a central role in the development of many behaviors, appropriate as well as inappropriate, and that its application in a number of studies has not been formally acknowledged. In what follows, I will describe three aspects of negative reinforcement as it relates to the applied situation: first, undesirable behavior acquired or maintained through negative reinforcement; second, the treatment of negatively reinforced behavior; and third, negative reinforcement as therapy. This organization departs somewhat from that used in reviews of the basic research literature and has been adopted here to highlight the relevance of particular issues to the applied researcher. Much of the research included here has been done with the developmentally disabled population because there is a high prevalence of significant behavioral disorders in this group and because it provides a narrow but adequate focus for discussion.

CURRENT CONCEPTUALIZATION OF NEGATIVE REINFORCEMENT

Before proceeding, it may be helpful to clarify terminology and to delineate the defining features of negative reinforcement. The purpose of this digression is to show that the task of determining whether a given contingency is an example of negative reinforcement may not always be a simple one. Although there has been little confusion regarding the effect of negative reinforcement, describing its operations has posed a challenge to many beyond the level of the beginning student.

The process of negative reinforcement typically involves the removal, reduction, postponement, or prevention of stimulation; these operations strengthen the response on which they are contingent (Hineline, 1977). Removal and reduction of ongoing stimulation typically produce behavior that
is called "escape," whereas postponement and prevention of stimulus presentation produce behavior that is called "avoidance." "Typically" is used as a qualifier throughout because the terms negative reinforcement, escape, and avoidance are subject to confusion under certain conditions, as the following will illustrate.

In commenting on the distinction between positive and negative reinforcement, Michael (1975) reviewed a number of historical points related to terminological usage. More important, he noted that some stimulus changes associated with an increase in behavior are difficult to classify as "presentation" (positive reinforcement) versus "removal" (negative reinforcement), and that the use of either description may be nothing more than an arbitrary and incomplete abbreviation for the static "prechange" and "postchange" stimulus conditions as well as for what transpires in between. For example, is a change in temperature more accurately characterized as the presentation of cold (heat) or the removal of heat (cold)? Problems such as this led Michael to suggest that "The distinction between two types of reinforcement [positive vs. negative], based in turn upon the distinction between presentation and removal simply can be dropped" (p. 44). An additional basis for distinguishing between positive and negative reinforcement was suggested first by Catania (1973) and later by Hineline (1984), who noted that "... if a stimulus or situation is to be reducible or removable by some response, that response must occur in its presence. In contrast, positively reinforced responses necessarily occur in the absence of the stimuli upon which reinforcement is based" (pp. 496–497). Such a distinction is not without its own problems, as can be seen in the previous example. Is responding prior to a temperature change more accurately described as responding in the presence versus the absence of heat (cold)? Another problem with this distinction is encountered when one considers the difference between escape (responding in the presence of stimulation), and avoidance (responding in the absence of stimulation), both of which are examples of negative reinforcement.

In many applied situations, it is possible to iden-

tify unambiguously a stimulus change as one involving presentation (e.g., of physical contact) or removal (e.g., of a token) and to determine whether or not the response of interest occurs in the presence or the absence of stimulation. However, because research outside of the laboratory is subject to greater variation in and less control over a multitude of potentially relevant stimuli, the motivational features of some stimulus changes are difficult to specify. Consider, as a case in point, Osborne's (1969) "Free-time as a reinforcer in the management of classroom behavior," which examined the out-of-seat behavior of six students. During the baseline condition, the students worked for approximately 4 hours daily without recess, and data showed that the target behavior occurred frequently. During treatment, students could earn 5 min of free time at the end of every 15-min work period by remaining in their seats, and data showed a decrease in out-of-seat. It is interesting to note the target behavior. Defined and recorded as "out-of-seat," free time was made available for its absence; this type of contingency usually is described as differential reinforcement of other behavior (DRO). However, the instructions given to students specified that they were to remain in their seats, suggesting "in-seat" as the functional target. If so, free time was made available for the occurrence of in-seat behavior; this type of contingency is not considered an example of DRO. Depending on how one characterizes "free time," (i.e., the availability of preferred activities vs. the termination of nonpreferred activities), we would label the contingency as one involving positive or negative reinforcement for in-seat behavior. Osborne suggested both of these possibilities in his discussion and perhaps this is why he did not place an adjective in front of the term "reinforcer" in the title of the article.

As a field, we have not attended carefully to the important distinction that Osborne drew. His study is regarded as a seminal piece of research in the applied literature for expanding our notion of what constitutes a reinforcing event and for demonstrating very nicely the effects of group contingencies, although the exact nature of the contingency is still unclear. A number of interesting replications and
extensions have appeared in JABA (e.g., Aaron & Bostow, 1978; Baer, Rowbury, & Baer, 1973; Barrish, Saunders, & Wolf, 1969; Harris & Sherman, 1973, 1974; Long & Williams, 1973; Maloney & Hopkins, 1973; Medland & Stachnik, 1973), but none have included further discussion or analysis of free time contingencies as positive versus negative reinforcement. Although such analyses may have little or no impact on outcome (i.e., in either case, behavior will have been increased), our general tendency to overlook a negative reinforcement interpretation may lead to undue emphasis on the numerous forms that free time may take at the expense of considering important features of the environment that free time replaces. That is, if free time serves as negative reinforcement, its only essential component may be alteration or termination of the preceding aversive situation.

In a more general sense, the complete analysis and specification of conditions in effect prior and subsequent to responding was the primary basis underlying Michael’s (1975) suggestion to eliminate the distinction between positive and negative reinforcement. It appears unlikely that the terms “positive” and “negative” will be deleted from our technical vocabulary in the near future; nevertheless, researchers should be cognizant of the fact that the two are potentially interchangeable and that failure to consider both possibilities may have a limiting effect on experimental procedure, interpretation, and subsequent application.

UNDESIRABLE BEHAVIOR ACQUIRED AND MAINTAINED BY NEGATIVE REINFORCEMENT

Hineline (1977) noted that a typical negative reinforcement paradigm includes three features: the presence of aversive stimulation, the availability of a response, and a suitable contingency between the response and the stimulation. Any behavior thus developed or maintained, including a variety of disruptive, destructive, aggressive, self-injurious, and otherwise problematic acts, could be considered “normal” or “adaptive” in that it is the orderly outcome of specific conditioning operations. The term “undesirable” is used here only as a means of classifying behaviors that are considered inappropriate given the usual social context.

An initial question of particular interest to those working in applied areas relates to factors that determine the form of the response. Acquisition of negatively reinforced behavior has been a subject of interest to basic researchers as well because it has been found that some topographies are more readily produced than others. For example, the treadle-press and shuttle responses of pigeons are more easily controlled by negative reinforcement than is the key peck, which is highly responsive to positive reinforcement (Ferrari, Todorov, & Graeff, 1973; Foree & LoLordo, 1970; MacPhail, 1968; Rachlin & Hineline, 1967; Smith & Keller, 1970). Similar data based on the study of different species have provided some support for the hypothesis that negative reinforcement involves selective control over preexisting “species-specific defense reactions” to aversive stimulation (Bolles, 1970, 1971). This account, however, does not provide adequate explanation for the wide range of human behaviors that apparently is susceptible to negative reinforcement. A more likely explanation is that aversive stimulation initially produces one or more of a variety of responses characteristic of both human and nonhuman subjects, including flinching, freezing, jumping, visual scanning, and related and diffuse motor activity (see reviews by Davis, 1979; Hutchinson, 1977; Myer, 1971), and that the eventual and more elaborate form of the behavior is determined by the individual’s previous history and the prevailing contingency.

Thus, many of the serious behavioral disorders that are seen in, for example, mentally retarded individuals may be a function of negative reinforcement applied to a particular behavioral repertoire and shaped over time. It is possible that certain instructional sequences (e.g., requests or even the appearance of specific training materials or the instructor) become discriminative for aversive stimulation in the form of physical contact, which is a common element in many teaching routines. At first, the stimulation and its associated cues may produce behaviors similar to those noted above. If,
NEGATIVE REINFORCEMENT

however, other behaviors have been successful in eliminating similar types of stimulation in the past, their eventual occurrence should not be surprising. Tantrums, attempts to flee, or destruction of materials are examples of such behavior, particularly if the individual is unskilled at more subtle or socially acceptable forms of escape. Although disruptive, these behaviors usually are not considered insurmountable barriers to instruction. A number of informal and formal interventions (e.g., proceeding in spite of the tantrum, “scooting” the individual’s chair under a table and backing both against a wall, bolting the materials to the table, etc.) are successful in managing disruptive behavior in some cases. In other cases, the interventions may provide a means for shaping more serious forms of escape. The immediate result of aggression for the individual toward whom it is directed suggests that physically harmful acts could serve as very effective escape behaviors, and their ability to terminate aversive instruction is most likely a function of the relative size and strength of client and trainer. Finally, self-injurious behavior, if severe enough, will quickly terminate any situation.

Data relevant to a negative reinforcement hypothesis for the development of behavior disorders in the mentally retarded exist in retrospective form only because it would be unethical to produce pathological behavior in humans when it does not already exist. Nevertheless, support for such a hypothesis can be found in several studies. Carr and Durand (1985) and Weeks and Gaylord-Ross (1981) showed that several different topographies of inappropriate behavior occurred more frequently during a “difficult task” condition when compared to an “easy task” condition, suggesting that the former condition contained aversive properties and that the resulting behavior was escape- or avoidance-motivated. Carr, Newsom, and Binkoff (1976) examined variables that apparently exerted stimulus control over the self-injurious behavior of a psychotic boy. In one of their experiments, they presented the boy with three alternating situations: a free-play period, a condition in which the experimenter spoke descriptive sentences to the child (e.g., “The sky is blue”)—this was called the “tact” condition, and a third condition in which the experimenter spoke instructions to the child—this was called the “mand” condition. Higher levels of self-injurious behavior were associated with the mand condition. Carr, Newsom, and Binkoff (1980) conducted a similar analysis of aggressive behavior in two boys, showing that aggression was more likely to occur when demands were present than when they were absent. Finally, Iwata, Dorsey, Slifer, Bauman, and Richman (1982) described a general methodology that allowed one to differentiate self-injury associated with positive versus negative reinforcement. In one of the conditions, self-injury was followed by adult attention; in another, self-injury produced brief escape from adult demands. Some subjects consistently exhibited self-injury during the latter condition, suggesting that their behavior was more sensitive to and maintained by negative reinforcement.

It is important for us to identify how environments that we create may provide negative reinforcement for undesirable behaviors. When faced with situations in which our students and clients are disruptive, we should immediately examine the antecedent as well as the consequent conditions to determine if the difference between the two provides reduction of aversive stimulation, keeping in mind that negative reinforcers may be just as idiosyncratic as positive ones. If we conclude that our clients and students exhibit bizarre and potentially dangerous behaviors to terminate instruction, we might question whether or not our well-intentioned efforts to teach are in our clients’ best interest; at the very least, we must question one or more aspects of our teaching technique. Perhaps most important from the standpoint of contingencies, our ability to identify negative reinforcement as a maintenance variable for undesirable behavior may directly influence treatment selection and outcome. This is particularly true with respect to extinction and time-out. Their use typically calls for one or more therapist responses (e.g., turning away from the client, removing stimuli from immediate access, removing the client from the setting, etc.) that terminate the ongoing situation. Studies conducted with nonhuman (Appel, 1963; Azrin, 1961; Thompson,
TREATMENT OF NEGATIVELY REINFORCED BEHAVIOR

A number of procedures based on the application of extinction, differential reinforcement, and punishment have been evaluated as treatments for problematic behavior of unspecified origin. Their use with behavior maintained by negative reinforcement will be discussed in this section, along with an additional procedure involving stimulus fading.

Extinction

Traditional time-out will not provide for the extinction of behavior that has been maintained by negative reinforcement, but other procedures might. One rather obvious possibility is elimination of the supposed aversive stimulation and its related cues, which should produce a reliable decrease in escape or avoidance behavior (Boren & Sidman, 1957; Shnidman, 1968). However, as Hineline (1977) has noted, this procedure may not be a true extinction operation. The complete removal of aversive stimulation during extinction of negatively reinforced behavior can be considered analogous to the continuous presence of, for example, food during extinction of positively reinforced behavior. Both procedures amount to noncontingent reinforcement, which removes the basis for responding and indirectly reduces the frequency of behavior. That is, if food is always present during extinction of food-maintained behavior, there is no basis for responding; a similar situation exists if shock is always absent during extinction of shock-avoidance behavior. Following these procedures, food removal or, alternatively, reappearance of the shock should immediately produce the target response (see Michael, 1982, for an extended discussion of this topic).

A more appropriate extinction procedure would entail continued presentation of the aversive stimulus or its cue and elimination of the consequence that was provided formerly (i.e., avoidance or escape). In this manner, the basis for responding (aversive stimulation) remains, but reinforcement does not (Bankart & Elliott, 1974; Coulson, Coulson, & Gardner, 1970; Davenport, Coger, & Specter, 1970; Schiff, Smith, & Prochaska, 1972). Techniques derived from this type of extinction actually have been used for a number of years in the treatment of clinical phobias and provide the major theoretical basis for interventions collectively known as “impllosion therapies” (Levis, 1979).

An example of extinction for negatively reinforced behavior was reported recently by Heidorn and Jensen (1984). After noting that demand-related situations were associated with an increase in their subject’s self-injurious behavior, a treatment was developed that included the following: (a) continued presentation of demands, (b) physical guidance to complete the requested performance contingent on the occurrence of self-injury, (c) termination of the session contingent on compliance, and (d) gradual increase in performance criteria across sessions. Positive reinforcement in the form of praise, food, and physical contact also was provided, but its role as an active component of treatment may have been minimal. A similar procedure was used in one of the experiments reported by Carr et al. (1980) on the treatment of aggression. Extinction consisted of belting the subject in a chair to prevent escape while a therapist wearing protective gear sat across a table from him. The intervention differed from that used by Heidorn and Jensen in that no attempt was made to deliver instructions during extinction sessions; instead, demands were introduced after aggressive behavior was eliminated almost completely.

As with extinction of positively reinforced behavior, it is possible to foresee situations in which extinction of negatively reinforced behavior might not be in the immediate best interest of either the client (as in the case of severe self-injury) or the
therapist (as in the case of aggression). Extinction procedures may be compromised further by the potential effects of what procedurally may resemble noncontingent aversive stimulation (see earlier discussion on acquisition of avoidance responding). To the extent that these "elicited" responses occur during the extinction of negatively reinforced behavior in applied situations, attempts to increase alternative behaviors, as well as to reduce the target behavior, may be disrupted. Finally, research showing that time-based delivery of aversive stimulation can maintain (Powell & Peck, 1969) and even increase (Kelleher, Riddle, & Cook, 1963; Sidman, Herrnstein, & Conrad, 1957) the rates of avoidance behavior suggests that schedule-related variables and the subject's previous history may be important considerations in the use of extinction.

Differential Reinforcement

Applications of reinforcement to decrease a target behavior (differential reinforcement of other behavior [DRO], differential reinforcement of incompatible behavior [DRI], etc.) are well documented in the applied literature, although the maintaining variable for the target behavior rarely is noted. The reinforcement contingency itself typically involves the use of positive reinforcement, and discussion here will be similarly confined. Applications of negative reinforcement will be addressed separately.

An experiment designed to examine the suppressive effects of differential reinforcement on negatively reinforced behavior may take several forms. First, access to an appetitive reinforcer (e.g., food) could be made contingent on the absence of the target behavior (DRO) while the escape/avoidance contingency is still operative. Although this approach might be considered unusual, it may resemble very closely situations in the natural environment in which DRO is implemented without attempting to identify the behavior-maintaining contingency. To my knowledge, this study has not been reported in the basic literature, probably due to difficulties associated with equating reinforcement. It is possible that this type of study has been reported in the applied literature but that it was not explicitly identified.

A second approach might involve appetitive reinforcement for a competing behavior (DRI) with the escape/avoidance contingency again operative. Ruddle, Bradshaw, Szabadi, and Foster (1982) studied human operant performance (button pressing) using exactly this procedure. They presented subjects with concurrent avoidance/positive reinforcement schedules, and obtained matched responding when the schedules were equated (this was made possible by using points exchangeable for money). Performance shifts were correlated with schedule shifts roughly in a manner predicted by Herrnstein's (1961) matching law. Our assessment research on self-injury (Iwata et al., 1982) provides an approximation to the Ruddle et al. methodology. During one condition, we presented to subjects a series of instructional demands. Compliance was followed by praise and physical contact from the experimenter, whereas the occurrence of self-injury produced a 30-s time-out. Data gathered during that study, as well as those collected since, indicate that both responses are likely to occur; in other words, positive reinforcement for compliance alone does not suppress avoidance-motivated self-injury. Another example of differential reinforcement was reported by Kelley, Jarvie, Middlebrook, McNeer, and Drabman (1984). They provided token reinforcement (stars) for reductions in the pain behavior (screaming, interfering, etc.) of two children undergoing open burn treatment. The procedure was moderately effective in that reductions in pain behavior averaged less than 50%. The findings of Iwata et al. and Kelley et al. are consistent with those of Ruddle et al., indicating that positively reinforced behavior competes with but does not suppress avoidance or escape responding that is reinforced concurrently. In contrast, Carr et al. (1980) were able to obtain almost complete elimination of aggression in one of their subjects by introducing positive reinforcement to an existing demand situation. They did note, however, that their second subject was not responsive to the positive reinforcement and that a different treatment (see previous discussion of extinction) was used.

A third experiment might examine reinforcement, as described in either of the above examples,
combined with extinction (continued presentation of aversive stimuli and prevention of escape). The Heidorn and Jensen (1984) study on self-injury, described previously, is an example of this approach. From an applied perspective, their procedures represent optimal treatment because contingencies were provided for the inappropriate as well as the appropriate behavior. However, here the effects of reinforcement are inseparable from those of extinction, and a clearer interpretation would require comparative analysis (reinforcement plus extinction vs. reinforcement alone vs. extinction alone).

The studies described here remain prototypical for the most part because very little research has been reported on the use of differential (positive) reinforcement with escape and avoidance behavior. On purely ethical grounds, and for the purposes of establishing and strengthening alternative behaviors, the use of positive reinforcement seems critically important. On the other hand, its therapeutic effects as primary treatment for negatively reinforced behavior have yet to be demonstrated. Based on the small amount of data available, one might expect that positive reinforcement is more likely to produce beneficial results if the negatively reinforced behavior is extinguished concurrently or if the density of positive reinforcement is noticeably greater than that of the negative reinforcement.

**Punishment**

Contingent aversive stimulation for negatively reinforced behavior is the functional complement of DRO for positively reinforced behavior, in that prevention of aversive stimulation (negative reinforcement) is contingent on the absence of responding. Procedural curiosities aside, we know very little about the effects of punishment on human escape and avoidance, in spite of the many applied studies on punishment published to date. For example, the literature on self-injury, in which most of the current applied research on punishment can be found, contains only two studies reporting the use of punishment for behavior described as avoidance motivated. One of the elements in the Heidorn and Jensen (1984) multiple-treatment approach consisted of physical guidance to complete a requested performance, contingent on the occurrence of self-injury. It is interesting to note that the particular form of stimulation used as punishment may have been exactly the same aversive stimulation whose prior removal served as negative reinforcement; if so, the treatment amounted to a perfect reversal of the maintaining contingency. Borreson (1980) also reported a case study of multiple treatment for avoidance-motivated self-injury; however, the punishing stimulus—"forced running" up and down a stairway—appeared to be unrelated to the prior function of the behavior.

Punishment of negatively reinforced behavior presents significant complexities not found with positively reinforced behavior because it involves aversive stimulation following responses for which such stimulation already plays an important role as an eliciting, discriminative, and motivating event (for extensive reviews of this topic, see Davis, 1979; Fowler, 1971; Hineline, 1981; and Morse & Kelleher, 1970, 1977). The major issues are summarized here. First, the eliciting properties of aversive stimulation, described previously with respect to acquisition and extinction, are relevant in the case of punishment. Although elicited behavior may not necessarily compromise the use of punishment, it may have a deleterious effect on the overall treatment program. Second, punishment with the same stimulus used during escape or avoidance training may acquire discriminative properties for responding as a result of reinstating the conditions under which escape originated, thereby occasioning the very behavior being punished. For example, several studies have shown response maintenance and even facilitation when shock-preventing behavior was followed by the presentation of shock (e.g., Appel, 1960; Sandler, Davidson, & Malagodi, 1966). Third, schedule-related variables can determine whether contingent stimulation serves as either punishment or reinforcement. Kelleher and Morse (1968) and McKearney (1972) found that responding developed as avoidance behavior was suppressed under dense schedules of punishment but facilitated under thinner schedules. Finally, it has been noted that punishment intensity and the pres-
ence or absence of avoidance contingencies may have an interactive effect on behavior. Sandler, Davidson, Greene, and Holzschuh (1966) imposed high-, intermediate-, and low-intensity shock as punishment for ongoing avoidance behavior and found greater response persistence under the high-intensity condition. However, when the avoidance contingency was later removed (i.e., responding produced shock but no longer prevented it), the high-intensity condition produced the most rapid response suppression.

The use of punishment should always be considered very carefully, and even greater caution should be taken when there is reason to believe that the target behavior has been maintained by negative reinforcement. Findings from the basic research literature suggest, although in a very tentative manner, that a stimulus different from that associated with prior avoidance should be used, that the schedule of punishment should be a continuous one, and that “mildly aversive” stimuli may produce greater response suppression than more intense stimulation when the prevailing avoidance contingency remains operative. On the other hand, data from the Heidorn and Jensen (1984) study indicated that, within the context of their multiple-treatment approach, the relationship between stimulation used as negative reinforcement and punishment may not be an important one.

**Stimulus Fading**

In contrast to approaches in which a contingency is directly manipulated, fading consists of altering one or more features of stimuli that occasion the target behavior. Various types of stimulus fading have been used for many years in the treatment of clinical fears and phobias, dating back to the work of Jones (1924). Contemporary formulations differ greatly along procedural dimensions (actual vs. representational stimulus presentation, the presence or absence of reinforcement and punishment) as well as on underlying theory (respondent vs. operant conditioning). The operant model of stimulus fading to reduce escape or avoidance behavior involves (a) initial identification of response-producing stimuli, (b) stimulus alteration to the point where responding does not occur, (c) presentation of the altered stimuli with a gradual return to their original state, and (d) extinction of escape behavior.

Approximations to the stimulus fading approach can be found in two studies previously discussed. Heidorn and Jensen (1984) decreased and then gradually increased the frequency with which response-producing stimuli (demands) were presented, although they did not withdraw them entirely at the beginning of treatment. In contrast, Carr et al. (1976) were able to reduce self-injury by embedding demands within entertaining stories, although the stories were never faded out nor were additional demands faded in. The results of both studies suggest that more complete evaluations of treatment based on fading are warranted. One potential advantage of fading over extinction and punishment might be the complete elimination of escape behavior from the outset of treatment.

**NEGATIVE REINFORCEMENT AS THERAPY**

I have noted previously that free-time contingencies might function as negative reinforcement, although that possibility has been seldom acknowledged. Free time also may be one of the few contingencies in the applied literature that represents pure escape in that the stimulation (work) is relatively continuous and can be reduced or terminated but not avoided. A great majority of applications make use of time- or trial-based presentation of stimuli preceded by cues, which produces avoidance behavior. Examples of negative reinforcement used to strengthen desirable behavior will be discussed in this section, grouped according to similarities in either procedure or problem.

**Behavioral Engineering**

The earliest examples of negative reinforcement to develop or maintain appropriate behavior published in *JABA* made use of apparatus-delivered stimulation. Azrin and his colleagues conducted two such studies. The first (Azrin, Rubin, O’Brien, Ayllon, & Roll, 1968) established automated measurement and control over postural slouching. An
apparatus built into a shoulder harness operated
closure of a circuit when slouching occurred. This
action produced an audible click, followed 3 s later
by a 55-db tone. The contingency was an ingenious
one in that it included aspects of both free-operant
and discriminated avoidance plus escape. Mainte-
nance of correct posture (free-operant) avoided the
click, postural correction (discriminated) during the
3 s following the click avoided the tone, and cor-
rection during the tone provided escape. Subjects
consisted of 25 adults, all of whom showed re-
ductions in slouching while the device was worn.
When the contingency was reversed for two of the
subjects, both showed increases in slouching. In a
later study, Azrin and Powell (1969) evaluated an
apparatus to increase pill taking in six subjects. The
pill dispenser produced a 50-db tone every 30 min,
an arbitrary between-pill interval. The tone could
be turned off by pushing a knob on the case, which
also delivered two pills. One might expect that this
arrangement would produce escape behavior ini-
tially, followed by free-operant avoidance, although
data to that effect were not presented. A third study
involving apparatus-delivered negative reinforce-
ment was conducted by Greene and Hoats (1969).
Their subject was an adult male assigned to a cor-
rectional unit whose task was to sort computer
cards, for which he earned cigarettes. He also was
allowed to watch TV while performing the task.
During the treatment condition, if the subject did
not complete a task cycle within a specified interval
of time (avoidance), visual and auditory output
from the TV were distorted and remained that way
until the work cycle was completed (escape).

Toilet Training and Incontinence

The presence or absence of elimination is more
than the occurrence or nonoccurrence of a response.
Sphincter contraction as well as relaxation is in-
volved, and negative reinforcement has been used
in the management of both behaviors. Hansen
(1979) incorporated escape and avoidance contin-
gencies in the treatment of nocturnal enuresis in
two children. When a device placed in the bed
detected urination, an apparatus located 4 ft away
produced a 70-db tone, which was followed in 7
s by a 95-db tone. By remaining dry, the child
could avoid the first tone; if the child urinated, he
or she could escape the first tone and avoid the
second by immediately getting out of bed and turn-
ing the unit off. O'Brien, Ross, and Christophersen
(1986) recently used negative reinforcement to pro-
duce elimination. As part of an overall treatment
program for four encopretic children, the authors
wanted to establish morning control over bowel
movements. To do so, parents had the children sit
on the toilet each morning for 5 min; failure to
defecate a minimum equivalent of one-fourth of a
cup during that time was followed immediately by
insertion of a suppository. A second administration
was given if the first did not produce the desired
outcome.

Overcorrection

Originally designed as a means for eliminating
accidents during toilet training (Azrin & Foxx,
1971), overcorrection consists of a group of tech-
niques whose common feature is repetitive perfor-
mance of motor activity. Overcorrection is one of
the most thoroughly studied and frequently used
methods for reducing the frequency of a wide range
of undesirable behaviors (see Foxx & Bechtel, 1983,
for a review). The procedures are considered to be
derivatives of punishment and are applied contin-
gent upon the occurrence of a target behavior. At
the same time, overcorrection can serve as negative
reinforcement in at least two ways. First, because
the procedure calls for performance of activities that
apparently are aversive, a therapist always is at hand
to ensure compliance through continued instruction
and, if necessary, physical guidance. Thus, the client
can avoid repeated instructions and potentially in-
trusive physical contact through continued perfor-
mance of the required activity. Avoidance behavior
is also produced when overcorrection is applied
contingent upon the absence of a desirable response.
For example, Foxx (1977) showed that 5 min of
functional movement training, involving the prac-
tice of varying head positions, was superior to food
and praise in developing and maintaining eye con-
tact in three retarded children. Examination of the
overcorrection literature yields a number of in-
NE GATIVE REINFORCEMENT

stances in which the procedure was used to increase rather than decrease behaviors, including class attendance (Foxx, 1976), repetitive tasks (Carey & Bucher, 1983), sharing (Barton & Osborne, 1978), speech (Matson, Esveldt-Dawson, & O’Connell, 1979), and spelling accuracy (Foxx & Jones, 1978; Ollendick, Matson, Esveldt-Dawson, & Shapiro, 1980).

Error Correction During Instruction

Rodgers and Iwata (1987) recently conducted a survey whose initial focus was on response prompting as an adjunct during behavioral acquisition. We quickly determined that most prompting occurs following an error and expanded the analysis to include all events that can be made contingent on incorrect responses. Negative reinforcement was not a subject of interest at the outset, but some of the techniques that were found suggest that it plays a much more prominent role in the instructional process than is currently acknowledged.

The most dramatic example is a study by Kircher, Pear, and Martin (1971) entitled “Shock as punishment” [emphasis added] in a picture-naming task with retarded children.” In one experiment, two children were exposed to the following two treatments: (a) token reinforcement for correct picture-naming responses, and (b) token reinforcement for correct responses plus shock for either errors or a response latency greater than 5 s. The shock condition produced superior results. Because the token reinforcement remained constant across the two conditions and because, regardless of how one defines the target behavior (i.e., errors vs. correct responses, inattention vs. attention), the desirable performance was a correct picture name, the procedure clearly represents an avoidance contingency in that correct responses made within 5 s of the cue prevented the delivery of the shock.

The Kircher et al. (1971) study represents a rather extreme use of negative reinforcement to increase desirable behavior, one that cannot be defended on ethical grounds today. However, less dramatic but analogous situations are quite common in the literature on instructional technology. It has become standard practice to follow errors with statements of disapproval (Dunlap & Johnson, 1985; Rincover & Newsom, 1985; Schreibman, 1975), physical guidance (Haring, 1985; Luyben, Funk, Morgan, Clark, & Delulio, 1986; Sprague & Horner, 1984), session-lengthening procedures consisting of either time-out (Barrera & Sulzzer-Azaroff, 1983; O’Brien & Azrin, 1972) or remedial learning trials (Nutter & Reid, 1978; Page, Iwata, & Neef, 1976; Richman, Reiss, Bauman, & Bailey, 1984), and so on. Thus, in addition to producing positive reinforcement in the form of experimenter praise, correct responses also may function to avoid aversive social and physical stimulation and to effectively reduce the duration of training sessions (this latter point is potentially significant, for it has been shown that complex setting events or stimulus situations, and not just discrete stimuli, can function as negative reinforcers [Krasnegor, Brady, & Findley, 1971], and that reduction of avoidance-session durations can itself serve as negative reinforcement [Mellitz, Hineline, Whitehouse, & Laurence, 1983]).

Behavioral Replacement Strategies

Given that aversive stimuli are ubiquitous and that escape is highly adaptive in their presence, it is usually the form, rather than the function, of escape and avoidance behavior that presents a problem. This raises the possibility of eliminating inappropriate forms of escape by negatively reinforcing appropriate alternatives; in essence, replacing one behavior with another but not eliminating the function of the original. The concept is rather straightforward in principle and is analogous to lever switching by nonhuman subjects following a change in reinforcement schedule (e.g., De Villiers, 1974).

Carr and Durand (1985) recently provided an example of this strategy by teaching three children, who had tantrums when faced with difficult tasks, how to request help from the teacher. When a child exhibited the appropriate response (“I don’t understand”), brief escape was provided in the form of teacher assistance.

Another example is drawn from our work on feeding disorders. While attempting to increase the
oral acceptance of food in four children, Riordan, Iwata, Finney, Wohl and Stanley (1984) found that one child, who was fed through a gastrostomy tube, did not respond well to positive reinforcement because her baseline rate of acceptance was virtually nonexistent. She also had resisted a number of forced feeding regimens; these practices were aversive and it appeared that her success in defeating them constituted negative reinforcement. Her treatment consisted of the following components: (a) the presentation of a redundant cue—"Take a bite"—immediately followed by (b) the presentation of food on a spoon. If acceptance of food did not occur within 3 s, (c) her mouth was held open and the food was deposited. This procedure thus resembled very closely a discriminated avoidance contingency in which one avoidance behavior (active food refusal) was replaced with another (opening the mouth and accepting food) by allowing it to prevent forced feeding.

Other Examples

In addition to the Kircher et al. (1971) study on academic performance described earlier, one can find instances in which negative reinforcement has been used—in a highly intrusive manner—to increase appropriate social behaviors. Lovaas, Schaeffer, and Simmons (1965) used escape and avoidance in training two autistic children to approach adults. Prior to treatment, the children frequently engaged in stereotypic behavior and showed no social responsivity or appropriate play. Treatment consisted of presenting the instruction, "Come here," followed by shock delivered through a floor grid. The shock was terminated when the child moved toward the therapist (escape). Both children quickly learned to approach the adult in response to the verbal instruction (avoidance). In defense of their use of electric-shock avoidance, the authors presented data indicating that increases in approach behavior were accompanied by more frequent displays of affection and decreases in stereotypy and aggression.

A less intrusive but similar procedure was used by Fichter, Wallace, Liberman, and Davis (1976) in an attempt to improve the social skills of a chronic and withdrawn schizophrenic male. The researchers targeted three aspects of his conversational behavior: voice loudness, duration of verbal responding, and keeping his hands on the armrests of his chair while speaking. A therapist approached the client, called his name, and asked him to converse about one of several predetermined topics. During treatment, the staff member continuously monitored the target behaviors; if any failed to meet criterion, the staff member would nag loudly one or more of the following: "Longer!," "Louder!," or "Put your hands on the armrests of the chair!," and would continue to nag at 3-s intervals until the target behavior occurred. Although effective, it should be noted that the contingencies used by Fichter et al. would not be considered typical consequences for social interaction; in fact, it is entirely possible that the appearance of the therapist would become discriminative for withdrawal. Data similar to those provided by Lovaas et al. (1965) showing that social behavior (e.g., conversations initiated by the subject or his response to approach by a therapist) improved outside of treatment sessions would have been informative with respect to this question. One can only assume that there was no generalized improvement, based on a comment made by the authors:

"... his [the subject's] last interaction before... [being discharged from the]... unit was to tell one staff member how much he disliked the unit and the staff" (Fichter et al., 1976, pp. 384–385).

SUMMARY

In this article I have attempted to point out a number of ways in which negative reinforcement is relevant to behavioral development and its subsequent modification in the applied situation. My review has been a selective one in that I made no attempt to summarize the large and varied literature on aversion, implosion, and desensitization therapies often used in the treatment of alcoholism, smoking, phobic reactions, sexual disorders, and related clinical problems. Still, the applied examples
represent a thorough cross section of research published in JABA over the past 20 years, and a number of general conclusions and implications can be drawn from the work described here.

Historically, applied analyses of behavior have failed to acknowledge escape and avoidance as potentially common and powerful sources of reinforcement. Evidence of this can be found in work on severe behavioral disorders such as self-injury, in which discussions of etiology have focused primarily on attentional factors rather than on those related to escape; in research on contingencies such as free time, in which free time as negative reinforcement has not been a subject of analysis; in studies in which avoidance contingencies have been inaccurately described as punishment; and in research on instructional processes, in which a variety of avoidance contingencies have been used but not evaluated or even described as such.

A second and more optimistic conclusion supportable by work described here is that an applied technology of negative reinforcement is emerging. The work is somewhat scattered at present and little is known in some areas. Nevertheless, under each of the topics included in the present discussion—behavioral development, treatment of negatively reinforced behavior, and therapeutic uses of negative reinforcement—research activity has increased in recent years, and we are beginning to see investigations of common procedures to the point where categorization is both possible and useful. This evidence of growing interest suggests that negative reinforcement may be one of the most significant areas of applied research during the coming years. Having made that prediction, what remains is to offer some prompts to help ensure its accuracy.

The area of behavioral development is particularly problematic because applied researchers often are faced with situations in which the behavior of interest has a long, complex, and unknown history. In fact, the most important difference between laboratory- and field-based research, at least from a behavioral standpoint, is the lack of control over history that is characteristic of applied research. In some cases (perhaps even most), behavioral history may be irrelevant if a sufficiently powerful contin-

egative can be found. In other cases, however, implementation of treatment without consideration of developmental factors, or treatment selection based on a consideration of topography alone, may produce a number of unnecessary failures and subsequently may limit our ability to determine the basis for differential outcome.

Laboratory researchers solve problems related to history by controlling its course in a naive animal. Although applied researchers rarely can exercise this option, they can make a unique contribution by developing methods for “unravelling” behavioral history, to the extent that it is possible. Our realization that behavioral development through negative reinforcement can produce the same topography as that resulting from positive reinforcement, time and time again, may provide the impetus for continued refinement in the analysis of behavioral function. For example, a number of researchers have concluded that unitary accounts of severe behavioral disorders in the developmentally disabled are unsatisfactory and have begun to establish methodologies for identifying the functional properties (one being negative reinforcement) of disorders such as pica (Mace & Knight, 1986), self-injury (Carr & Durand, 1985; Iwata et al., 1982), and stereotypy (Durand & Carr, 1987). Most recently, Bailey (1987) has proposed the term “behavioral diagnostics” to describe a general strategy for isolating the bases of problematic behavior. Continued work in this area and extension of the relevant methodologies to other human problems are essential if we are to develop a mature technology of behavior. In the meantime, researchers should be encouraged (perhaps by editors) to seek out and include more detail on subjects’ behavioral histories. In addition to the usual demographics offered (e.g., age, sex, grade or functioning level, etc.), which provide little information relevant to a behavioral analysis, it would be helpful to provide some account of factors related to behavioral development and maintenance. As evidence supporting a negative reinforcement interpretation accumulates, we will be increasingly compelled to formalize our anecdotal observations and to confirm these observations through manipulation.
Because applied research often is concerned with a problem as it actually exists, the treatment of negatively reinforced behavior will provide perhaps the greatest opportunity for creative work. For example, research on the extinction of positively reinforced behavior has included variation and extension (e.g., time-out, exclusion, seclusion, contingent observation, time-out ribbon, movement suppression), parametric analysis (e.g., duration, delay, schedule, changeover requirement), and comparison (e.g., with differential reinforcement, response cost, and punishment). None of these questions have been addressed adequately in applied research on the extinction of negatively reinforced behavior, and a similar situation exists with respect to differential reinforcement and punishment.

Research on the treatment of negatively reinforced behavior will require consideration of issues that are different than those relevant to the treatment of positively reinforced behavior. These issues have been noted previously, and some have been the focus of laboratory research for several years. Stimulus selection, schedules, and intensity, for example, may differentially affect the outcome of contingent aversive stimulation for ongoing avoidance behavior. Therefore, it will be important for applied researchers to become acquainted with basic findings on negative reinforcement. As a result, we may find that methodologies and procedures developed in the laboratory can be extended to the applied situation so as to facilitate analysis and treatment.

Research on the use of negative reinforcement may take several interesting directions. First, negative reinforcement may provide an alternative means for establishing behavior when attempts to use positive reinforcement fail (e.g., as in the case of Riordan et al., 1984, in which a child's operant level of eating was nonexistent). If so, we will want to know the behaviors for which specific contingencies are useful and the conditions under which they should be applied. Second, it appears that the acquisition of adaptive behavior in our training programs is at least partially a function of negative reinforcement. Future research must evaluate the roles of escape and avoidance within the training context so that (a) we will have a proper estimate of the effectiveness of commonly used positive reinforcers (the results of this estimate may indicate that more potent reinforcers are needed), (b) we can determine whether procedures such as remedial trials, physical assistance, and so on, serve any useful function and if that function is one of negative reinforcement, and (c) we can base future training successes on the planned rather than the accidental use of negative reinforcement. A third promising application involves further elaboration of behavioral replacement strategies. If we are willing to entertain the assumption that it is impossible to eliminate all sources of aversive stimulation, the use of such stimulation to alter the topography of escape and avoidance behavior, from an undesirable one to a tolerable one, makes eminent sense from a clinical standpoint.

A final cautionary note. Some of the applied research included in this review was selected specifically to show that negative reinforcement can form the basis of highly intrusive intervention. In at least one sense, negative reinforcement might be considered more intrusive than punishment because, with negative reinforcement, presentation of the aversive stimulus is contingent on the absence, rather than the occurrence, of behavior. Therefore, as with punishment, we should conduct research on negative reinforcement with great care and under the appropriate conditions to determine how it might be used effectively and humanely, its limitations, and its proper role within the larger realm of currently available treatment.

REFERENCES


Fox, R. M. (1976). Increasing a mildly retarded woman's


Michael, J. (1975). Positive and negative reinforcement: A distinction that is no longer necessary; Or a better way to talk about bad things. Behaviorism, 3, 33–44.


Received August 25, 1987
Initial editorial decision August 30, 1987
Revision received September 14, 1987
Final acceptance September 15, 1987
Action Editor, Jon S. Bailey