Theoretical Implications of the Neurotic Paradox as a Problem in Behavior Theory: An Experimental Resolution

Thomas G. Stampfl
University of Wisconsin-Milwaukee

Why do human phobias last for months or years when such behavior should undergo extinction? This failure of extinction or persistence of self-defeating behavior of human disorders was labeled by Mowrer as the neurotic paradox. The paradox is cited by an ever-increasing number of critics who challenge any laboratory-based learning model of human psychopathology. Laboratory research, of course, omits essential requirements in the analysis of behavior, and the principles derived from such analyses must be combined in order to explain complex human behavior. Validation for a behavioral model can thus be achieved if (a) basic principles inferred from observation of humans treated with a laboratory-derived extinction procedure (e.g., implosive therapy) are combined with (b) principles examined in laboratory research that are combined to generate unique predictions that correspond to known features of human phobic behavior. The latter evidence is briefly reviewed in research demonstrating sustained responding over one thousand consecutive active avoidance responses with complete avoidance of the "phobic" CS for an initial single shock trial. Differential reinforcement for responses to early sequential stimuli depends on minimal work requirement, and reinforcement by timeout from avoidance. This combination of factors effectively precludes extinction to main conditioned aversive stimuli for nonhumans, as it does for human phobias. Support for a laboratory model of human phobia is thereby attained.

Laboratory-derived principles of learning and conditioning have played an important role in the development of significant applied and conceptual advances in the etiology and treatment of human psychopathological conditions (Dollard & Miller, 1950; Kalish, 1981; Mowrer, 1939; Skinner, 1953, 1954; Stampfl & Levis, 1967; Ullmann & Krasner, 1965). The interplay between reinforcement and extinction principles culled from an enormous amount of research in the learning laboratory proved remarkably fruitful in accounting for many facets of maladaptive behavior as seen in neurotic and psychotic patients. Recently, however, there has been a rising insistence in

the behavior therapy literature that central difficulties exist in learning formulations directed to the problems of human psychopathology (Barlow, 1982; Boulougouris, 1982; Eysenck, 1976, 1979; Rachman, 1976, 1977; Wilson, 1982).

The neurotic paradox is often presented as one of the major limitations upon which *any* laboratory-derived system of behavior founders. An example of the challenge to behavior theory based on the neurotic paradox is that of Eysenck (1976):

Skinner, his colleagues and pupils have in recent years advanced a theory which is offered as a general theory of human conduct; this is based essentially on operant behaviour and the effects of positive and negative reinforcements. . . . Summaries of recent attempts to apply these principles to neurotic and psychotic behaviour are interesting in that they fail to deal with the central paradox of neurosis, as defined by Mowrer; as he indicates, it is precisely because neurotic behaviour does not follow Thorndike's law of effect, or Skinner's laws of reinforcement, that a special theory is required to account for these departures. . . . Restatement of the general theory does not solve the riddle. (p. 252)

THE NEUROTIC PARADOX

In 1926, Freud (1926/1936) described a basic theoretical problem, the resolu-

This article is a greatly expanded version of a symposium presentation conducted at the meeting of the American Psychological Association, Washington, DC, August, 1986: Learning research and the neurotic paradox, in Steven C. Hayes (Chair). *Stop running on empty: Some clinical implications of basic research*. The author is indebted to Patricia D'Auria for running some of the animals, and for help in the design of the proper control conditions for some of the research reported in the paper.

Reprints may be obtained from Thomas G. Stampfl, Department of Psychology, University of Wisconsin-Milwaukee, Milwaukee, WI 53201.

tion of which defied repeated attempts at solution:

We consider it entirely normal that a little girl should weep bitterly at the age of four if her doll is broken, at the age of six if her teacher reprimands her, at the age of sixteen if her sweetheart neglects her, at the age of twenty-five, perhaps, if she buries her child. Each of these grief-occasioning situations has its proper time and vanishes with its passing; but the later and more definite ones remain operative throughout life. We should be rather surprised, in fact, if this girl, after she had become a wife and mother, should weep over some knickknack getting broken. Yet this is how neurotics behave. Although in their mental apparatus there have long since developed all the agencies necessary for dealing with a wide range of stimuli, although they are mature enough to be able to gratify the greater part of their needs themselves . . . they nevertheless behave as though the old danger situation still existed, they remain under the spell of all the old causes of anxiety. . . . But how does this situation come about? Why are not all neuroses merely episodes in the individual's development which become a closed chapter when the next stage of development is reached? Whence comes the element of permanency in these reactions to danger? . . . In other words, we find ourselves abruptly confronted once again by the oft-repeated riddle: What is the source of neurosis, what is its ultimate, its specific, underlying principle? After decades of analytic effort this problem rises up before us, as untouched as at the beginning. (pp. 89-92)

Mowrer (1948, 1950, 1952) labeled the phenomena described by Freud the "neurotic paradox." He perceived the paradox as posing severe difficulties for learning-based explanations of neurotic behavior; "It is the question as to why so-called neurotic behavior is at one and the same time *self-defeating and yet self-perpetuating*, instead of self-eliminating" (1950, p. 351). Mowrer regarded the paradox as absolutely central for clinical theory and practice. He argued that since laboratory learning studies repeatedly demonstrate that fears extinguish with repetition of the stimulus or situation in the absence of the unconditioned aversive stimulus, "Why is it, then, that in neurosis we have fears which appear to have long outlived any real justification but which stubbornly persist or which may even augment to the point of seriously incapacitating the individual?" (1952, p. 679)

Hull (1929) discussed the essence of the paradox in his description of the di-

lemma of the conditioned defense reaction in terms of Pavlovian conditioning and extinction. He noted that in the case of defense (avoidance) reactions made by organisms in their natural environment to incidental stimuli correlated with genuine injuries that (1) Pavlovian extinction would doom them to suffer genuine injury repeatedly or (2) Pavlovian extinction was ineffective for genuine injury of a serious nature, and that the organism was doomed to respond indefinitely to irrelevant or incidental stimuli present at the time of the original injury. Hull speculated that human phobia might be understood in terms of the second alternative. However, neither alternative seemed to be especially adaptive biologically, and they held little appeal for Hull. He concluded, "The problem presents a fascinating field for experimental investigation." (p. 511)

Amsel and Rashotte (1984) observed that Hull's formulation cast the "dilemma" into an avoidance model of learning, and that the problem Hull identified in 1929 continues to the present as a central difficulty for virtually all modern theorists interested in developing complete models of avoidance learning:

Hull posed a dilemma for conditioned-reflex mechanisms in the form of a problem that theorists have confronted ever since. . . . by definition, once successful avoidance has been conditioned, the process of experimental extinction should set in and avoidance behavior should deteriorate. (p. 27)

It is beyond the scope of this paper to treat in detail the ramifications contained in the vast avoidance literature as to the "deterioration" of the avoidance response as it is seen in laboratory research. Unfortunately, the literature contains so many inaccuracies that clarification of the "facts" alone would require a separate paper. One thing seems certain, however: If human psychopathology is explained in terms of an avoidance model as seen, for example, in human phobias, some variables or mechanisms are required to account for the "failure of extinction" to occur. Human phobics continue to manifest avoidance behavior for months and years. If this behavior is interpreted as avoidance responding made to aversive

or fear evoking stimuli as seen in the laboratory, it clearly appears to represent a failure of extinction. In Eglash's (1952) words:

In our attempt to cope with this paradox of self-perpetuating, self-defeating behavior, which Mowrer has called the crucial challenge to personality and learning theory, we stand today where Freud stood, with the problem unsolved. (p. 378; see also Seeman, 1951)

B. F. SKINNER

A remarkable characteristic of Skinner's development of the experimental analysis of behavior was his willingness to extrapolate knowledge gained from this analysis to a wide variety of significant human behaviors. Of course, much of the extrapolation was, more or less, on the speculative side, with no direct experimental verification of the validity of the translation. Skinner (1953, 1954) dealt with numerous clinical phenomena such as both traditional psychodynamic therapy and nonanalytic therapy. He included detailed descriptions of various psychodynamic concepts such as repression and the various defense mechanisms. Among the many human phenomena that were interpreted in terms of a behavioral analysis were clinical phenomena such as sibling rivalry, dreams, drug addiction, hallucinations, masochism, anger, depression, and delusions of grandeur. Skinner embellished his descriptions of behavioral principles applied to clinical phenomena with numerous examples of common human behavior which skillfully illustrated the points he was making. Looked at in terms of the main thesis of the present paper, such applications made over 30 years ago appear as nothing less than astonishing.

It was inevitable, however, that such widespread application to "complex" human behavior understood with principles derived mainly from animal research would result in some serious difficulties. Domjan (1987) recently expressed the contemporary point of view when he said:

The belief that results of animal research can (and should) be used to understand human behavior was

most ardently pursued by B. F. Skinner and his students, who created a worldview of human behavior based on animal research (see, for example, Skinner, 1953). Such enthusiastic embrace of animal research has been discouraged by increasing evidence of the complexity of various forms of human behavior. (pp. 558-559)

In the area of human psychopathology it certainly appears that the behaviors involved are complex. But, in fact, a persistent examination of clinical phenomena progressively gained over a period of many years reveals that the complexity is vastly overrated. In brief, the most complex psychopathological human behaviors reduce to relatively simple behavioral principles. Unfortunately, it sometimes takes an incredibly long time to recognize the combination of behavioral principles that provides a clear and testable solution to the difficulties that appear to exist.

SKINNER ON THERAPY

For Skinner, therapy could proceed in two essentially different directions. First, therapy could proceed according to the behavior analytic view of how traditional psychotherapy operated, with its emphasis on information contained in the personal conditioning history of the patient. Second, therapy could focus on the direct manipulation of variables known to change behavior, such as the systematic use of positive and negative reinforcement, without undue reliance on details of the personal conditioning history. Of course, this is an oversimplification, since both alternatives might ideally be used for maximal therapeutic impact, as Skinner himself suggested (1953, p. 380). Of interest, however, is that the first alternative was almost exclusively based on extinction of aversive stimuli *symptomatically pivotal to unique features of the conditioning history*, whereas the second alternative focused on direct manipulation of controlling variables without undue emphasis on extensive use of the conditioning history.

For reasons not entirely clear (see Stampfl & Levis, 1967), behavior therapy has almost exclusively utilized the second alternative. Perhaps Skinner's de-

scriptions of traditional psychotherapy as "rich in explanatory fictions" (1953, pp. 372-373) taken to be the determinants of neurotic behavior, or his comment about the "'effects of anxiety' . . . The middle term [anxiety] is of no functional significance, either in a theoretical analysis or in the practical control of behavior" (1953, p. 181), or his emphasis on "the use of positive reinforcement in the practical control of behavior" (1953, p. 384) led many behavior therapists to deemphasize the first alternative.

Skinner, however, clearly delineated a description of avoidance behavior that was pregnant in its implications for the way psychotherapeutic treatment of human psychopathological behavior might proceed. Treatment would consist of procedures that permitted the deconditioning of aversive stimuli that control symptomatic behavior. The controlling aversive stimuli were seen by Skinner to result from the conditioning history of the patient. Skinner's explanation of traditional psychotherapy was couched explicitly in terms of his behavioral analysis of avoidance behavior, stated essentially as an escape from conditioned aversive stimuli: "So far as we are concerned here, it [psychotherapy] may be described simply in this way; the therapist constitutes himself a nonpunishing audience" (Skinner, 1953, p. 370), that, in effect, facilitates through extinction the report of additional punishing (aversive) events by the patient. Skinner (1953) described the consequences of this therapeutic relationship as follows:

The appearance of previously punished behavior in the presence of a nonpunishing audience makes possible the extinction of some of the effects of punishment. This is the principal result of such therapy. Stimuli which are automatically generated by the patient's own behavior become less and less aversive and less and less likely to generate emotional reactions. The patient feels less wrong, less guilty, or less sinful. As a direct consequence he is less likely to exhibit the various forms of operant behavior which, as we have seen, provide escape from such self-generated stimulation. (p. 371)

Skinner (1953) clearly specifies how these maladaptive behaviors emerge,

What is "wrong" with the individual who displays these by-products of punishment is easily stated. A

particular personal history has produced an organism whose behavior is disadvantageous or dangerous. (p. 372)

By accounting for a given example of disadvantageous behavior in terms of a personal history and by altering or supplementing that history as a form of therapy, we are considering the very variables to which the traditional theorist must ultimately turn for an explanation of his supposed inner causes. (p. 379)

That extinction plays a direct function as a treatment manipulation is seen in Skinner's suggestion for various phobic-like fears,

Eliminating a conditioned response is also a common practical problem. For example, we may want to reduce the fear reactions which have come to be evoked by people, animals, air raids, or military combat. Following the procedures in the conditioned-reflex experiment, we present a conditioned stimulus while omitting the reinforcing stimulus responsible for its effect. (p. 58)

It is interesting that Skinner (1953) also included "anxiety" as a possibly more important stimulus condition than some of his other statements concerning anxiety would imply when he said:

Although the emotional aspect of anxiety may be distinguished from the conditioned aversive effect responsible for avoidance behavior, it is possible that the emotion is also aversive. Avoidance responses may be interpreted as in part an escape from the emotional components of anxiety. Thus we avoid the dentist's office, not only because it precedes painful stimulation and is therefore a negative reinforcer, but because, having preceded such stimulation, it arouses a complex emotional condition which is also aversive. (p. 179)

To sum up Skinner's main concern with traditional theoretical conceptions of the causes of human psychopathology, one must conclude that the concern is not necessarily with the environmental events specified by such theorists, but by the use made of them in their metaphysical constructions. In Skinner's (1954) critique of psychoanalytic theory, he specified his main objection to such metaphysical constructions:

For example, early punishment of sexual behavior is an observable fact that undoubtedly leaves behind a changed organism. But when this change is represented as a state of conscious or unconscious anxiety or guilt, specific details of the punishment are lost. When, in turn, some unusual characteristic of the sexual behavior of the adult individual is

related to the supposed guilt, *many specific features of the relationship may be missed that would have been obvious if the same features of behavior had been related to the punishing episode.* [italics added] (p. 231)

VALIDATION OF THE LABORATORY MODEL

The main purpose of this paper is to illustrate a form of validation for a laboratory model of human phobic behavior. To some degree, of course, validation for a theoretical avoidance model is conferred to the extent that current treatments derived from such a model are effective in reducing psychopathological behavior. Implosive therapy, as developed by Stampfl (1961), is one such example. It was based on Mowrer's two-factor theory of avoidance learning (see London, 1964), but its theoretical underpinnings deviate in no significant way from the first alternative delineated by Skinner (1953, 1954; see Levis, 1980; Stampfl, 1966, 1970; Stampfl & Levis, 1967, 1969, 1973, 1976). The therapy has been effective in treating a wide range of human psychopathological disorders (see Boudewyns & Shipley, 1983; Levis & Hare, 1977, for summaries of experimental outcome studies with clinical patients; see Hirt & Greenfield, 1979, for the efficacy of more general avoidance models).

Treatment success, however, does not constitute complete validation of the theoretical model that predicted it. Alternative interpretations of such successes are always possible and they sometimes appear intuitively more plausible than the laboratory model utilized to generate the new techniques in the first place. This is especially so for those clinicians who lack experimental training in behavioral science. In discussing research on a narrow experimental problem, Dinsmoor (1983) summarizes well my views about the use of behavioral principles postulated to validate a laboratory model of human psychopathology:

There is a strong tendency in contemporary writings, from the research report to the undergraduate text, to follow the fashion of the moment, to abandon—almost frivolously—well-established principles in favor of whatever fits the spirit of the time.

This, however, is not the way to construct a strong and orderly science, one capable of bending nature to its will. (p. 704)

Dinsmoor was reacting primarily to the use of cognitive explanations for laboratory phenomena. The approach described below utilizes basic behavioral principles in developing another source of validation for an avoidance model of human psychopathology without cognition. To simplify the exposition, the presentation is restricted to that of human phobia.

LIMITATIONS: CS AVOIDANCE AND ONE TRIAL LEARNING

Mineka (1985) has reviewed and summarized a number of major limitations to avoidance models of human phobia. Since avoidance models have been based to a large extent on animal studies, she correctly states that "the crucial criteria on which the model should be evaluated are whether the known features of conditioned avoidance responses seen in the laboratory resemble the known features of naturally occurring human phobias" (p. 204). Two of the central limitations she specifies are, first, that human phobics

generally go to great lengths to avoid their phobic stimulus (the presumed CS in an avoidance paradigm), whereas animals trained to avoid aversive stimulation in the laboratory are trained to avoid the US, not the CS. In fact, in spite of a number of attempts in the literature to train animals to avoid CSs (for aversive USs) there have apparently been no successful attempts to do so. . . . We see here a significant and important difference in a crucial feature of avoidance learning as studied in the laboratory and in naturally occurring phobias. (p. 204)

And, second,

that most phobias acquired through traumatic conditioning are acquired in one trial, whereas only a very few laboratory studies have ever demonstrated one trial fear conditioning. . . . (p. 206)

Both of these limitations may be challenged. Avoidance of CSs in serial CS laboratory studies has been clearly demonstrated (Boyd & Levis, 1976; Kostanek & Sawrey, 1965; Levis, 1979; Levis & Boyd, 1979). In fact, Mineka, herself, cites some of these studies in her 1985 article, stating that

well-trained avoidance responses can be markedly resistant to extinction, and the persistence can be potentiated by the use of serial CSs. . . . In fact, with serial CSs animals are known to make short latency responses to early portions of the stimulus and thereby manage to avoid later portions of the CS [these are different CSs] which elicit higher levels of anxiety (p. 226)

Even Sidman avoidance schedules can produce relative avoidance of the CS (warning stimulus) when the work component (response rate) correlated with a specific response-shock interval is introduced (Sidman, 1957; Sidman & Boren, 1957). To determine the aversiveness of the warning stimulus (CS) based on its proximity to shock, the work component must be experimentally controlled. When it is, relative avoidance of the warning stimulus (CS) emerges (Logan & Boice, 1968). The major reference used by Mineka and others to support the contention that animals do not avoid the CS is that of Sidman (1955a) in which the work variable was not controlled (see Mineka, 1979a, 1979b). Sometimes individual animals will develop strong avoidance of the CS even when the work component clearly favors waiting for the warning stimulus. Sidman (1955b, Experiment 1-B, p. 217) reported a cat that made 1,502 successive lever-press avoidance responses in a three and one-half hour session without receiving a single shock. He stated that the cat made 246 responses to the warning stimulus (tone), and therefore one can presume that it made 1,256 responses *before* the warning stimulus. Sidman schedules, however, represent a complex analysis (see Hinesline, 1977, 1981) that requires a separate article to examine how they relate to the specific thesis of this paper (see, e.g., Field & Boren, 1963). Nevertheless, it seems reasonable to argue on the basis of conditioned aversive or even conditioned fear theories of avoidance behavior that stimuli correlated with work would themselves acquire aversive or fear eliciting properties. Thus, the various properties of Sidman avoidance schedules that include warning stimuli may be interpreted as conditioned aversive stimuli that the animal will escape from if given the opportunity to do so. If darkness were

correlated with a higher rate of responding (work) than the rate of responding with light, then darkness as a stimulus will be more aversive than the light stimulus. The relative aversiveness will be reversed if the response-shock interval controlling response rate (work requirement) is also reversed for the two stimulus conditions.

An examination of "one trial" fear conditioning in human phobics also reveals that a strong argument can be made that multiple trials are actually involved in the "single trial" human case, as "trial" is defined in the laboratory.

For the moment, however, let us accept Mineka's (1985) limitations at face value. Acquisition of the conditioned fear or aversiveness to the phobic stimulus CS should occur in one or two trials. To conform more fully to the human phobic's behavior, in fact, some additional requirements are necessary. Human phobics frequently avoid the phobic situation almost completely. They do this repeatedly without any clear effects of extinction—the neurotic paradox. Relative avoidance of the CS therefore is not acceptable in the case of the human phobic. Virtually 100% avoidance of the conditioned aversive "phobic" stimulus is required. A more compelling laboratory test of the alleged parallel existing between human phobia and a laboratory-based animal avoidance model, such as Mowrer's or Skinner's, requires the following three characteristics: (1) avoidance of the CS (phobic stimuli) should occur on almost all trials; (2) acquisition of fear or aversiveness to the CS (phobic stimuli) should occur in one or two trials; and (3) the consequences of (1) and (2) should generate a very large number of successive (consecutive) active avoidance responses with few, if any, signs of extinction—again, the neurotic paradox.

One trial fear conditioning therefore should generate active avoidance responses that approach the persistence of avoidance seen in the human phobic and display the property of almost always avoiding the CS originally conditioned. Williams and Watson (1985) describe such avoidance of a CS:

Height phobics typically refrain from taking part in skiing long before they are standing in line for the chair lift; social phobics decline invitations to future social events from the safety of their living rooms; and bridge phobics plan vacation routes precisely to insure that they will not be confronted with an intimidating span. (pp. 138–139)

It should be noted that if a basic behavioral principle can be identified that would produce almost complete avoidance of the conditioned aversive CS, it would solve the neurotic paradox since the conditions for extinction would be absent. Hilgard and Marquis (1940) summarized Guthrie's position concerning the problem:

Guthrie's (1938) theory of neurotic symptoms is based largely on the proposition that neurotic behavior never gets extinguished because the neurotic individual avoids the stimuli in the presence of which extinction might take place. Thus fear of high places can never be overcome if the person with that fear always avoids high places. (p. 293)

Allport (1937) supported his functional autonomy explanation of persistent human behavior with the description of single event acquisition of psychopathological disorders, "An experience associated only once with a bereavement, an accident, or a battle, may become the center of a permanent phobia or complex, not in the least dependent on a recurrence of the original shock" (p. 148). That a single experience may result in powerful aversive consequences in the human case is also mentioned by Skinner (1953):

Since conditioning may take place as the result of one pairing of stimuli, a single aversive event may bring a condition of anxiety under the control of incidental stimuli. The sudden death of a close friend, for example, is sometimes followed by a sustained depression which may be verbalized as a feeling that "something is going to happen," as a "feeling of doom," and so on. It is hard to deal with such a case. (p. 179)

In brief, a rigorous test illustrating the correspondence between a laboratory-generated model of avoidance behavior and the known features of human phobias would require 100% active avoidance responses based on a few seconds of electric shock in a single trial, resulting in virtually complete avoidance of the conditioned aversive "phobic" stimulus.

This is to say that the less stringent "trials to extinction criterion" typically used in avoidance studies employing clearly defined acquisition and extinction phases is not acceptable. Such extinction criteria permit the organism to obtain credit for a trial in resistance to extinction when, in fact, no avoidance response is made on that trial. If the extinction criterion is set at five consecutive failures to avoid, then a single avoidance response in any set of five trials permits the organism to obtain five trials in resistance to extinction. This is not the way human phobics behave.

A final comment is required concerning the frequent statements made by critics of laboratory-based behavior models of human behavior disorders. Experimental psychologists are typically interested in important problems not immediately related to human psychopathology. Laboratory research is ordinarily greatly simplified for its own purposes. Such simplification leads to the omission of requirements needed to demonstrate adequately the phenomena critics have asserted are necessary to validate a laboratory-based model of human psychopathology. Under such circumstances, it is not surprising that laboratory research has failed to demonstrate the relevant phenomena. In the words of Skinner (1953):

In a science of behavior we begin in the simplest way. We study relatively simple organisms with relatively simple histories and under relatively simple environmental conditions. In this way we obtain the degree of rigor necessary for a scientific analysis. Our data are as uniform and reproducible as, say, the data of modern biology. It is true that the simplicity is to some extent artificial. We do not often find anything like it outside the laboratory—especially in the field of human behavior, which is of primary interest. As a result those who are impatient to get on to bigger issues are inclined to object to the "oversimplified" formulations of the laboratory. Their objections take the form, as in the example of Boyle's Law, of setting forth apparent exceptions to the rule. Such criticism is useful if it points to facts which have been unseen or ignored. But very often the exceptions are only apparent; the existing formulation is capable of giving a good account of them when properly applied. . . . A common source of misunderstanding is the neglect of what happens when variables are combined in different ways. (pp. 204–205)

Quite clearly, many of the theoretical difficulties posed by critics might be explicable by a propitious combination of basic behavioral principles.

The development of a "laboratory" oriented treatment strategy based on extinction principles—implosive therapy—in the late 1950s led to some interesting observations concerning phenomena related to the symptomatic behavior of human patients (Stampfl, 1961, 1966). Many of these observations depended on a minute examination of the specific details underlying the characteristics of the conditioning history, as well as the specificity of details related to the aversive stimuli controlling the patient's behavior. This was, of course, precisely what Skinner (1954) had emphasized. Phenomena were viewed exclusively in terms of basic behavioral principles. The basis for combining variables in order to account for certain characteristics of complex human disorders became discernible. How could validation for these observations be achieved? It seemed reasonable to test the implications of these observations directly in the animal laboratory with unique predictions related to specific problems. The experiment described in what follows illustrates just one instance of a prediction for animals derived from observations of humans treated by implosive therapy. Unique predictive validity for the laboratory avoidance model and the exposure treatment technique of implosive therapy would thereby be attained, an exciting development, indeed (Krasner, 1970).

As early as 1959, it was evident that the aversive or anxiety evoking stimuli in the human patient were multiple, complex, and usually organized sequentially, and that contextual stimuli were important contributors to the total amount of fear or aversiveness controlling the symptomatic behavior of the patient. Avoidance (escape) behavior made to sets of stimuli early in sequence precluded exposure to stimulus subsets later in the sequence. Skinner (1953) provides several examples of human avoidance behavior that implies a sequential ordering

of different stimulus segments, each of which must be aversive to some degree. Clearly, such an arrangement would retard extinction since avoidance to early stimulus segments does not permit extinction to later stimulus segments in the sequence. The persistence of human symptomatic behavior (e.g., phobia) is partially explained by the serial or sequential ordering of stimuli. The conditions required for rapid acquisition of avoidance behavior could readily be inferred from an analysis of context stimuli and stimulus complexity, but a nagging problem remained. Although rapid acquisition and greatly increased resistance to extinction was attained with rats using serial stimuli, the absolute number of responses before extinction did not seem sufficient to parallel the resistance to extinction seen in the human patient. Rats would make responses to stimuli early in the sequence, but extinction, followed by exposure to subsequent stimuli, would take place. What might be described as secondary intermittent reinforcement slowed the process, but progressive exposure and extinction of the avoidance response was inevitable.

The problem seemed related to the fact that human phobics typically respond predominantly to stimuli early in the sequence, but rats did not display this strong "fixation" to early stimuli. Avoidance of the main source of aversiveness correlated with subsequent stimulus segments in the human phobic precludes extinction from taking place to the subsequent aversive stimulus segments, exactly as Guthrie (1938) described. But what behavioral principle accounts for this strong "back-up" to early stimuli? It is easy, of course, to declare that humans are essentially different from rats, and that the behavioral principles governing humans are qualitatively different from rats.

A closer examination of the "back-up" tendency in human phobics, however, revealed that additional reinforcing events were present when avoidance (escape) responses were made to stimuli early in the sequence: Human phobic behavior involves not only a set of complex stimuli

that vary in a sequential continuum, but the avoidance responses of the phobic are also variable and they differ in response cost such as increased work and other aversive consequences (Stampfl, 1983a). This variability is in contrast to the usual laboratory avoidance experiment where the response requirement is relatively constant. The animal runs or jumps or presses the lever with essentially the same response requirement on each trial. The work requirement or response cost for the laboratory animal is essentially the same whether latencies are short or long, and is true also of serial CS preparations whether the response is made early or late in the sequence of stimuli. "Short latency" responses made by human phobics, however, frequently reduce the response cost such as increased work.

In terms of the examples of Williams and Watson (1985), when a social phobic declines an invitation from the safety of the living room, such an avoidance response minimizes the amount of work involved. When a bridge phobic plans a vacation route so as to insure avoidance of bridges, it is evident that much less response cost is present than driving without a plan, encountering a bridge that is highly aversive, and only then stopping to plan a route where bridges are circumvented. Bridge phobics may on occasion fail to plan their route in advance. When this happens, however, they are punished by two consequences: (1) exposure to conditioned aversive stimuli (anxiety or fear) evoked by bridge stimuli and (2) the punishing response cost of work and other negative consequences correlated with the failure to plan the route in advance. In the human condition, such failure might include punishing consequences such as loss of time and money, social embarrassment, frustration, etc., as well as more work. It should also be evident that the planned vacation route may be made at any point along a spatial-temporal distance dimension to the bridge. This basic sequence is applicable to many if not the majority of human phobias.

In this view, the avoidance response typically "backs up" quite rapidly to

stimuli early in the sequence for the human phobic by the punishing consequences of increased exposure to more aversive stimulus segments and of increased response cost in the greater work requirement. Responses to early stimulus segments in the sequence minimize both sources of aversiveness. It is the summation of the two sources that produces the total amount of aversiveness in the situation.

Another reinforcing factor is present in the human situation. When the bridge phobic plans a route in advance or the social phobic declines an invitation to a future social event from the safety of a living room, the response is followed for a period of time by freedom from having to make avoidance responses. Timeout from avoidance responding (e.g., Baron, DeWaard, & Lipson, 1977) provides optimal conditions for reducing to a minimum both conditioned aversive stimuli and response cost.

Timeout from avoidance responding may provide another independent source of reinforcement for early responses by permitting self-generated positively reinforced behavior to be emitted during the timeout period. This function of timeout in the human situation appears obvious since it permits, for a period of time, a large repertoire of positively reinforced responses common to everyday human behavior. A similar interpretation is possible for the rat also, because timeout provides the opportunity to engage in uninterrupted grooming and other potentially self-generated positively reinforcing behaviors during the timeout period. It should also be noted that the human phobic's early responses followed by timeout maximizes timeout periods per unit of time.

There are other variables clearly operative in human phobic and other psychopathological behavior. The description given above, however, is sufficient to describe the main features relevant to the research presented below (Stampfl, 1983b, 1985, 1986). The experiment combines behavioral variables to generate behavior in rats that corresponds to

known features of human phobic behavior: One trial fear conditioning, complete avoidance of the phobic CS, and one thousand active avoidance responses with few, if any, indications of extinction.

COMBINING BEHAVIORAL PRINCIPLES

The strategy is to arrange avoidance responding for the rat to a more powerful serial CS that involves contextual changes along a spatial-temporal dimension. Avoidance responses made to stimuli early in the sequence will require less work than to stimuli later in the sequence. Timeout from avoidance will follow most immediately responses made to early stimuli in the sequence. Stimuli present during the timeout period are most distant from the aversive "phobic" CS complex originally conditioned. Timeout stimuli are minimally aversive. Secondary intermittent reinforcement effects are maximized with the arrangement of stimuli.

In an apparatus designed to permit rapid learning of the avoidance response, rats are trained to avoid a "phobic" stimulus complex along a spatial-temporal stimulus continuum where the response requirement is variable. That is, responses made to stimuli early in the sequence of stimuli (most distant from the "phobic" stimulus complex) require less work as defined by the distance traversed than to stimuli later in the sequence. Responses made to stimuli early in the sequence provide more timeout from avoidance per session than for responses made to stimuli later in the sequence. One shock trial of brief duration (dimensions of the grid are designed to permit rapid escape from shock) for all animals is delivered, and the current is disconnected. Therefore, all responses do not avoid the US. There is no shock US available after the initial shock trial.

Apparatus

Rats are trained in a straight alley runway (5 feet long, 12 in wide and 14 in high with no ceiling (1,524 mm long $\times$ 305 mm wide $\times$ 356 mm high). The

straight alley has a smooth floor. Added to the end of the 5 foot long straight alley is a black compartment 8.3 in deep, 12 in wide, and 14 in high (211 mm deep $\times$ 305 mm wide $\times$ 356 mm high) with a grid floor resting over a drop pan. The black compartment is covered with a transparent Plexiglas ceiling unlike the straight alley which has no ceiling. A removable black plywood barrier is used to separate the straight alley with the smooth floor from the black compartment with the grid floor. The side walls of the straight alley are white except for a 16 in (406 mm) black sidewall section immediately preceding the black shock compartment. A unique feature of the apparatus is that the smooth floor of the straight alley forms the top of a small remote-controlled commercial conveyor belt that can move the floor in either direction. Floor speed is set at 1.54 in (39 mm) per second when activated. (Rats can run the entire length of the alley in less than two seconds.)

Procedure

Rats receive 10 min of exploration time in the entire apparatus (5 min with the floor stationary and 5 min with the floor moving). All rats highly prefer the black compartment as indicated by time spent in it during this 10 min period.

Following this exploratory period, the rat is placed at the far end (photocell end) of the straight alley and permitted to return to the black compartment. The black barrier is placed in front of the black compartment, and 10 sec later footshock is delivered. Simultaneously, the black barrier is removed permitting escape from the black compartment. Rats reliably run to the far end of the straight alley where they remain motionless. Three minutes after escaping shock, the rat is moved in the direction of the black compartment by initiating floor movement in the direction of the black compartment. Rats typically remain motionless on the first trial until they enter the black sidewall section (16 in) that precedes entry to the black compartment. At this point, they promptly run to the far end of the straight

alley breaking the photocell and stopping the floor. Three minutes later a second trial is given by moving the floor, the response of running to the far end of the alley and breaking the photocell defines the active avoidance response. A 2 h session is given following this procedure. In the second session, a fixed-ratio for responding is gradually imposed until the rat makes 10 responses (breaks the photocell 10 times) for each 3 min timeout from avoidance responding. Two hour sessions are given several times a week with the FR 10 requirement.

Results

Following the single shock trial, rats reliably run to the far end of the straight alley where they remain motionless during the 3 min timeout period. On the first avoidance trial, rats remain motionless until the floor carries them to the black sidewall section where they respond quickly by running back to the photocell white end of the straight alley. Thereafter, rats quickly learn to make responses that minimize the distance traversed to break the photocell and obtain timeout from making avoidance responses. Rather than wait and be carried by the moving floor towards the "phobic" black compartment, rats quickly learn to make responses that minimize conditioned aversive stimuli, and the work requirement. They move a few inches rather than several feet. Under these conditions rats make more than one thousand successive avoidance responses over five 2 h sessions without a single entry into the "phobic" black compartment. In fact, there is little indication that reentry to the highly aversive "phobic" black compartment is imminent. Responses are predominantly made at the far end of the white alley and very infrequently is the rat carried to the black sidewall section preceding the black compartment before responding. Various controls indicate that the work component is independently aversive, and that the contingency between responding by breaking the photocell and obtaining timeout is critical to sustained responding. Rats given a single

shock trial and trained noncontingently with the same floor movement time, the same number of timeouts, and the same number of photocell clicks, in the same order as an animal contingently trained, extinguishes quite rapidly (usually fewer than 100 responses) by reentering the black grid compartment. The apparatus permits a number of manipulations that are not possible in the usual active avoidance apparatus. By using barriers that rest on the moving floor, the latency of active avoidance responses can be controlled and the point in the alley where they occur is under the control of the experimenter and not determined by the rat's avoidance behavior. Response delay or prevention trials of different durations can be introduced or shock trials given at any time.

DISCUSSION

Animal laboratory research typically does not include the requirements to demonstrate adequately the correspondence with known features of human psychopathological behavior such as phobias and avoidance learning. One example is that of the work requirement that differs in humans, but is maintained as a constant in laboratory studies of avoidance. By combining principles of behavior derived from an analysis of key variables present in human phobic behavior, it is possible to meet major objections to conventional models of avoidance learning that postulate simple conditioned aversive or fear evoking contingencies as controlling variables in such behavior. The present research provides evidence for one trial fear conditioning, complete avoidance of the "phobic" CS, and truly persistent avoidance responding that offers a solution to the neurotic paradox. The determination of the combination of variables responsible for the prediction were drawn from observations employing an animal model for the treatment of human disorders. Such a procedure confers strong additional validity to the animal model originally used for the treatment of humans. It is the case that the critical variables identified, such as the

sequential ordering of stimuli, secondary intermittent reinforcement, contribution of context stimuli, work requirement, timeout, and the critical function of the "backup" phenomenon have been less thoroughly examined than other variables in laboratory research. The use of variables less well investigated may add to, rather than subtract from, the strength of the present analysis. The predictions are not trivial, but represent solutions to central limitations that have been advanced to deemphasize the value of learning models that were successful in the development of new human treatment strategies in the first place. It was reasonable to specify these limitations because decades of research in the laboratory failed to reveal the phenomena thought critical to the human condition. By no means does the present analysis exhaust the possibilities for the solution of other "limitations" specified by critics. Also, there are many implications of the present research for a number of contemporary versions of avoidance learning.

In a word, the contemporary learning and conditioning literature provides marvelous opportunities for the further understanding of human clinical phenomena.

REFERENCES

- Allport, G. W. (1937). The functional autonomy of motives. *American Journal of Psychology*, 50, 141-156.
- Amsel, A., & Rashotte, M. E. (1984). *Mechanisms of adaptive behavior*. New York: Columbia University Press.
- Barlow, D. (1982). The context of learning in behaviour therapy. In J. C. Boulougouris (Ed.), *Learning theory approaches to psychiatry* (pp. 75-84). New York: Wiley.
- Baron, A., DeWaard, R. J., & Lipson, J. (1977). Increased reinforcement when timeout from avoidance includes access to a safe place. *Journal of the Experimental Analysis of Behavior*, 27, 479-494.
- Boudewyns, P. A., & Shipley, R. H. (1983). *Flooding and implosive therapy*. New York: Plenum.
- Boulougouris, J. C. (Ed.). (1982). *Learning theory approaches to psychiatry*. New York: Wiley.
- Boyd, T. L., & Levis, D. J. (1976). The effects of single component extinction of a three-component serial CS on resistance to extinction of the conditioned avoidance response. *Learning and Motivation*, 7, 517-531.
- Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, 6, 693-728.
- Dollard, J., & Miller, N. E. (1950). *Personality and psychotherapy*. New York: McGraw-Hill.
- Domjan, M. (1987). Animal learning comes of age. *American Psychologist*, 42, 556-564.
- Egash, A. (1952). The dilemma of fear as a motivating force. *Psychological Review*, 59, 376-379.
- Eysenck, H. J. (1976). The learning theory model of neurosis—A new approach. *Behaviour Research and Therapy*, 14, 251-267.
- Eysenck, H. J. (1979). The conditioning model of neurosis. *Behavioral and Brain Sciences*, 2, 155-199.
- Field, G. E., & Boren, J. J. (1963). An adjusting avoidance procedure with multiple auditory and visual warning stimuli. *Journal of the Experimental Analysis of Behavior*, 6, 537-543.
- Freud, S. (1936). *The problem of anxiety*. (H. A. Bunker, Trans.). New York: Norton. (Original work published 1926.)
- Guthrie, E. R. (1938). *The psychology of human conflict*. New York: Harper.
- Hilgard, E. R., & Marquis, D. G. (1940). *Conditioning and learning*. New York: Appleton-Century-Crofts.
- Hineline, P. N. (1977). Negative reinforcement and avoidance. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 364-414). Englewood Cliffs, NJ: Prentice-Hall.
- Hineline, P. N. (1981). The several roles of stimuli in negative reinforcement. In P. Harzem & M. D. Zeiler (Eds.), *Predictability, correlation, and contiguity* (pp. 203-246). New York: Wiley.
- Hirt, M., & Greenfield, H. (1979). Implosive therapy treatment of heroin addicts during methadone detoxification. *Journal of Consulting and Clinical Psychology*, 47, 982-983.
- Hull, C. L. (1929). A functional interpretation of the conditioned reflex. *Psychological Review*, 36, 498-511.
- Kalish, H. I. (1981). *From behavioral science to behavior modification*. New York: McGraw-Hill.
- Kostanek, D. J., & Sawrey, J. M. (1965). Acquisition and extinction of shuttlebox avoidance with complex stimuli. *Psychonomic Science*, 3, 369-370.
- Krasner, L. (1970). Comment. In D. J. Levis (Ed.), *Learning approaches to therapeutic behavior change* (pp. 205-207). Chicago: Aldine.
- Levis, D. J. (1979). The infrahuman avoidance model of symptom maintenance and implosive therapy. In J. D. Keehn (Ed.), *Psychopathology in animals: Research and clinical implications* (pp. 257-277). New York: Academic Press.
- Levis, D. J. (1980). Implementing the technique of implosive therapy. In A. Goldstein & E. B. Foa (Eds.), *Handbook of behavioral interventions: A clinical guide* (pp. 92-151). New York: Wiley.
- Levis, D. J., & Boyd, T. L. (1979). Symptom maintenance: An infrahuman analysis and extension of the conservation of anxiety principle. *Journal of Abnormal Psychology*, 88, 107-120.
- Levis, D. J., & Hare, N. (1977). A review of the theoretical rationale and empirical support for the extinction approach of implosive (flooding)

- therapy. In M. Hersen, R. Eisler, & P. Miller (Eds.), *Progress in behavior modification* (Vol. 4). New York: Academic Press.
- London, P. (1964). *The modes and morals of psychotherapy*. New York: Holt.
- Logan, F. A., & Boice, R. (1968). Avoidance of a warning stimulus. *Psychonomic Science*, 13, 53–54.
- Mineka, S. (1979a). The role of fear in theories of avoidance learning, flooding, and extinction. *Psychological Bulletin*, 86, 985–1010.
- Mineka, S. (1979b). New perspectives on conditioning models and incubation theory: A review of H. Eysenck's "The conditioning model of neurosis." *Behavioral and Brain Sciences*, 2, 178.
- Mineka, S. (1985). Animal models of anxiety-based disorders: Their usefulness and limitations. In A. H. Tuma & J. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 199–244). Hillsdale, NJ: Erlbaum.
- Mowrer, O. H. (1939). A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychological Review*, 46, 553–565.
- Mowrer, O. H. (1948). Learning theory and the neurotic paradox. *American Journal of Orthopsychiatry*, 18, 571–610.
- Mowrer, O. H. (1950). *Learning theory and personality dynamics: Selected papers*. New York: Ronald Press.
- Mowrer, O. H. (1952). Learning theory and the neurotic fallacy. *American Journal of Orthopsychiatry*, 22, 679–689.
- Rachman, S. (1969). Treatment by prolonged exposure to high intensity stimulation. *Behaviour Research and Therapy*, 7, 295–302.
- Rachman, S. (1976). The passing of the two-stage theory of fear and avoidance: Fresh possibilities. *Behaviour Research and Therapy*, 14, 125–131.
- Rachman, S. (1977). The conditioning theory of fear-acquisition: A critical examination. *Behaviour Research and Therapy*, 15, 375–387.
- Seeman, W. (1951). On a stimulus-response analysis of insight in psychotherapy. *Psychological Review*, 58, 302–305.
- Sidman, M. (1955a). Some properties of the warning stimulus in avoidance behavior. *Journal of Comparative and Physiological Psychology*, 48, 444–450.
- Sidman, M. (1955b). On the persistence of avoidance behavior. *Journal of Abnormal and Social Psychology*, 50, 217–220.
- Sidman, M. (1957). Conditioned reinforcing and aversive stimuli in an avoidance situation. *Transactions of the New York Academy of Sciences*, 19, 534–544.
- Sidman, M., & Boren, J. J. (1957). The relative aversiveness of warning signal and shock in an avoidance situation. *Journal of Abnormal and Social Psychology*, 55, 339–344.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Skinner, B. F. (1954.) Critique of psychoanalytic concepts and theories. In T. Millon (Ed.), (1967), *Theories of psychopathology* (pp. 228–235). Philadelphia: Saunders. (Reprinted from *Scientific Monthly*, 79, 300–305.)
- Stampfl, T. G. (1961, May). *Implosive therapy: A learning theory derived psychodynamic therapeutic technique*. Paper presented at the University of Illinois, Champaign, IL.
- Stampfl, T. G. (1966). Implosive therapy: The theory, the subhuman analogue, the strategy, and the technique. Part I: The theory. In S. G. Armitage (Ed.), *Behavior modification techniques in the treatment of emotional disorders* (pp. 12–21). Battle Creek, MI: Veterans Administration Publications.
- Stampfl, T. G. (1970). Implosive therapy: An emphasis on covert stimulation. In D. J. Levis (Ed.), *Learning approaches to therapeutic behavior change* (pp. 182–204). Chicago: Aldine.
- Stampfl, T. G. (1983a). Exposure treatment for psychiatrists? [Review of *Learning theory approaches to psychiatry*.] *Contemporary Psychology*, 28, 527–529.
- Stampfl, T. G. (1983b, November). *Truly persistent active avoidance responding with one shock trial*. Address of the President of the Hull-Spence Society. San Diego: CA.
- Stampfl, T. G. (1985, January). *Towards an experimental resolution of the neurotic paradox: The role of sequential stimuli, work requirement, and timeout from avoidance*. Paper presented at a meeting of the Winter Conference on Animal Learning, Winter Park, CO.
- Stampfl, T. G. (1986, August). Learning research and the neurotic paradox. In S. C. Hayes (Chair), *Stop running on empty: Some clinical implications of basic research*. Symposium conducted at the meeting of the American Psychological Association, Washington, DC.
- Stampfl, T. G., & Levis, D. J. (1967). Essentials of implosive therapy: A learning-theory-based psychodynamic behavioral therapy. *Journal of Abnormal Psychology*, 72, 496–503.
- Stampfl, T. G., & Levis, D. J. (1969). Learning theory: An aid to dynamic therapeutic practice. In L. D. Eron & R. Callahan (Eds.), *Relationship of theory to practice in psychotherapy* (pp. 85–114). Chicago: Aldine.
- Stampfl, T. G., & Levis, D. J. (1973). *Implosive therapy: Theory and technique*. Morristown, NJ: General Learning Press.
- Stampfl, T. G., & Levis, D. J. (1976). Implosive therapy: A behavioral therapy? In J. T. Spence, R. C. Carson, and J. W. Thibaut (Eds.), *Behavioral approaches to therapy* (pp. 189–110). Morristown, NJ: General Learning Press.
- Ullmann, L. P., & Krasner L. (Eds.), (1965). *Case studies in behavior modification*. New York: Holt.
- Williams, S. L., & Watson, N. (1985). Perceived danger and perceived self-efficacy as cognitive determinants of acrophobic behavior. *Behavior Therapy*, 16, 136–146.
- Wilson, G. T. (1982). The relationship of learning theories to the behavioural therapies: Problems, prospects, and preferences. In J. C. Boulougouris (Ed.), *Learning approaches to psychiatry* (pp. 33–56). New York: Wiley.