

Learning theory and the evolutionary analogy

Marion Blue

Erindale College, University of Toronto

1979

Keywords: learning, operant conditioning, classical conditioning, matching law, cognitive learning theory

Abstract

In this article, past comparisons of learning and evolution as analogous processes are discussed and some inaccuracies and omissions in those discussions are pointed out. The evolutionary analogy is examined for its ability to suggest solutions to five fundamental theoretical issues about learning - superstitions, why a reinforcer has the effect it does, the relationship among various procedures yielding learning, the relevance of the matching law to the problem of what reinforces an avoidance response, and whether behavioral and cognitive views of learning can be reconciled. In each case it is argued that the analogy is instructive.

1. Introduction
2. Vestiges and superstitions
3. Why does a reinforcer have the effect it does?
4. A single learning process
 - 4.1 The evolutionary analogue of discriminative control
 - 4.2 The evolutionary analogue of sensitization and habituation
 - 4.3 The evolutionary analogue of classical conditioning and conditioned suppression
 - 4.4 Summary
 - 4.5 Some evidence and implications
5. Act classes, the matching law and avoidance
6. A cognitive-reinforcement theory
 - 6.1 Blind variation and selection
 - 6.2 Place learning and regulative development
7. Conclusion

1. Introduction

The topic of this paper is the oft-drawn analogy between learning and evolution, between the selection of acts by reinforcement in learning and the selection of organisms by natural selection in evolution (Campbell, 1960; Gilbert, 1970, 1972; Pringle, 1951; Russel, 1962; Skinner, 1966; Staddon, 1975; Staddon and Simmelhag, 1971). The analogies drawn in the past have not always been precisely accurate or complete. When they are drawn in a more accurate and complete fashion they can be instructive. The evolutionary analogy suggests for example an

answer of a sort to the question of why a reinforcer has the effect it does. Perhaps an act becomes more probable when reinforced simply because it is adapted to or “competent” with respect to the reinforcer involved. Reinforcers in the classical position (i.e. initially preceding the act they reinforce) may be analogous to some environmental influence on phenotypes in evolution, inducing the development of organisms in an adapted direction, thereby altering future probabilities of occurrence. This possibility suggests the existence of a single learning process of which sensitization, classical conditioning, and instrumental and operant conditioning (including their negative counterparts - habituation, conditioned suppression and punishment) are aspects without reducing any of these forms of learning to any other. Herrnstein’s matching law is shown to be analogous to the most fundamental law of evolution, the Hardy-Weinberg equilibrium principle of population genetics theory. The implication derived, that avoidance may be a matter of the suppression, by punishment or induction, of competing behaviors can explain why the teaching of arbitrarily chosen avoidance responses is so difficult. Natural selection alters the relative frequency, not of actual morphologies, but of the genotypes which give rise to them. If reinforcement and punishment also select among programmes giving rise to act topographies instead of the actual topographies themselves, learning can be seen as both cognitive and a consequence of reinforcement and punishment. The cognitive view of learning as it stands today however is shown to be analogous to an inadequate Lamarckian rather than to a Darwinian theory. After a brief discussion of superstitions, each of these points will be taken up in turn.

2. Vestiges and superstitions

The phenomenon of learned superstitions (in which reinforcement, simply applied at

fixed intervals independent of behavior leads to the learning of an idiosyncratic response) has played a prominent role in discussions of the learning/evolution analogy. Skinner (1966) has interpreted superstitious acts analogous to vestiges in evolution. This is not a precisely accurate analogy. A vestigial analogy would exist if a feature which composed part of an act resulting in reinforcement, but one which did not itself appear to contribute to that result, was incorporated into the act learned anyway. By contrast, with a superstition, an entire act which just happens to precede reinforcement learned. In their discussion of the analogy, Staddon and Simmelhag (1971) have implied that phenomena analogous to superstitions are an inconceivable result of selection in evolution. It

would be like taking a population of white mice, breeding them for 20 generations without further selection for color, and then attributing the resulting white population to the results of 'accidental selection.' (p. 21)

Is the analogy of superstitious reinforcement so impossible to imagine in evolution? It may or may not be a very significant phenomenon there, but it is certainly not theoretically impossible. Many superficially plausible analogies exist in addition to vestiges - drift, correlative evolution and artificial selection for example. None of these are precisely correct however. Superstitiously learned acts tend to be seen as "accidentally" reinforced because we know that many other acts of the organism, if they had occurred at the right time, could just as easily have been learned instead. The reinforcer just happened to occur at the right stage of this one act's "life cycle." Imagine than a cyclically occurring selection pressure in nature, say the presence of some particularly useful next building material available only cyclically because it comprised part of the life cycle of some plant. Imagine two varieties of birds, both of which, at the right stage of their life cycle, could equally well make use of the material, for both of which it could

constitute a positive selection pressure therefore. But what if the life cycle of only one of these varieties coincided with that of the plant in such a way as to enable it to make use of this material? Obviously this variety would be selected for relative to the other. Nor is the “drift” often shown by superstitious acts, as through time one often gradually replaces another surprising. Since the only edge this act has over many other acts is its timing, and since the effect of the reinforcer is to make the act more probable and therefore on the average to make it occur slightly earlier, the time may come when the act happens just early enough for another act to occur after it, but prior to the reinforce. The process may begin again as this new act now becomes the one to be superstitiously reinforced and the potential for a perpetual drift from one to another is evident. On the other hand, this result is not inevitable. To the extent that variation exists within one of these act populations with respect to life cycle length and phasing, to the extent that these are dimensions on which members of the population vary, then fixation is possible.

The analogy reveals however that the term “superstition” suggesting “accidental” is somewhat of a misnomer. In both the evolutionary and the learning case, the item selected for has a real advantage over others - its life cycle length and its phasing. It is just that this advantage is of a somewhat different kind than the more familiar morphological or topographical kinds. In summary, learned behaviors almost certainly include vestigial components and evolution almost certainly includes phenomena analogous to superstitions, but the two are not the same.

None of the foregoing should be taken to imply that the fact that some stimulus event happens to follow some act is sufficient to explain an increase in the probability of occurrence of

that act, nor particularly, that Skinner's superstition explanation adequately deals with interim activities (Staddon and Simmelhag, 1971). The former problem of why a reinforcer has the effect it does may be clarified by asking the analogous question and examining its answer in evolutionary theory. The latter problem of interim activities leads deeply into the problem the influence of stimuli which precede rather than follow immediately observable behavior and what the evolutionary analogy can suggest about the role of such stimuli. Let us consider both of these issues in turn then.

3. Why does a reinforcer have the effect it does?

A common strategy in attempting to answer this question has been to seek a definition of "reinforcing events" which is both universal and operational and which has typically referred to the organism as a whole - for example Thorndike's "satisfying" (1911) or Hull's "drive-reducing" (1943). Such a strategy sought to render the law of effect testable. Rather than repeat here the litany of theoretical and empirical difficulties this strategy and others proposed by Meehl (1950) and Premack (1959, 1963, 1965) for example have encountered, let us consider the analogous question in evolutionary biology.

There, it is organisms which are analogous to acts in learning theory and there a positive selection pressure increases the frequency of occurrence of an organism of some type. Why does it have this effect? The answer employs a concept "adapted" or "fit" as a property of the organism involved. From the vantage point of the analogy with evolutionary theory, learning theory is missing a property ascribable to acts not to the whole organism, analogous to the term "adapted" or "fit." To fill this gap in learning theory, I suggest that we employ the adjective

“competent.” In evolutionary biology the quality of an organism of being adapted or fit must always be referred to a specific environment. The state of being adapted exists because of a relationship between the structure of an organism and features of its environment. There is no universal operational measure of the degree of adaptedness of an organism independent of its environment. Similarly, there is no universal operational measure of the direction or extent of selection that will be exerted by a particular environmental feature independent of the structure of the organism involved. I would suggest that the competence-reinforcement relationship in learning is similar to this adaptation-natural selection relationship in evolutionary biology. A competent act is one which in the circumstances present possesses features which make it likely to be reinforced. The state of being competent exists because of a relationship between the topography of an act and its environment, and there exists no universal measure of competence and reinforcement which are independent of each other. One way of conceptualizing the varying competence of acts in neurophysiological terms would be at least in part, as a matter of location. An act, comprising as it does the firing of a specific group of neurons, has a neurophysiological “location” even though that location may be widely distributed in space in the nervous system. However it is concretely that subsequent events alter the future firing probabilities of these neurons, it is possible that the stimulus events comprising a reinforcing state of affairs reach some locations but not others. Two acts differing to topography or neurophysiological location would vary then in their competence vis-à-vis a particular reinforcer, and two reinforcers would vary in the presence or absence of, and in the extent of, their reinforcing effect on the same act. The answer proposed then to the problem of reinforcement is that, just as selection depends upon the adaptedness or fitness of organisms, reinforcement may depend upon the competence of acts.

Compared with some previous attempts to solve the basis of reinforcement problem, it should be noted that the “competence” solution refers to a different unit of analysis. It refers to acts not to organisms. It is a competent act not a competent organism that is reinforced. In addition, it provides a universal but non-operational specification of reinforcing events. To say that whether an event will act as a reinforcer depends upon the competence of the act involved appears to be an analytic (or tautological) rather than a synthetic (or testable) statement. It is now widely recognized however that the distinction between these two classes of statements in scientific theories is far from being absolute (Kaplan, 1964, pp. 100-103; Suppe, 1977). The introduction of concepts into scientific theories is permissible on the condition that they have systematic import (e.g. contribute to a systematically unified theory of phenomena) and be empirically justified (Hempel, 1966, pp. 91-97). In this paper the concept of competence (coupled with other systematic extensions of the evolutionary analogy) will be used to formulate a single theory of what have heretofore been widely considered to be three distinct learning processes. Part of its justification therefore is its contribution to theoretical integration.

The analogous problem in evolutionary theory (adaptation and natural selection with the single measure being changes in gene or organism relative frequency) has led to the occasional suggestion that the adapted or fit term be dropped from the theory and only natural selection be retained. Natural selection just is differential reproduction according to this view rather than a force or cause which explains something. “It is only in the older and looser Darwinian sense that selection is genuinely a mechanism in nature” (Greene, 1958, p. 122). Note that this is similar to the approach Skinner has taken to the problem of reinforcement (Catania, 1973, 1978, 1979; Skinner, 1935). The concept of a reinforcer is not an unspecified and therefore tautologically

postulated cause of behavior change or learning according to this view, rather,

The contemporary operant view of reinforcement, is not explanatory at all. ... reinforcement is a name for a particular type of behavioral phenomenon. When we have shown that a response has become more likely because the response has produced a particular consequence, we say that the response has been reinforced. (Catania, 1978, p. 55)

A careful reading of the synthetic theorists of evolution shows that they do not support this view of the evolutionary case. In the most precisely formulated version of evolutionary theory, in population genetics theory, it is either readily stated (Mettler and Gregg, 1969, pp. 92-93), or implicitly acknowledged in the derivation of fitness and selection coefficients (Crow and Kimura, 1970, chap. 1 and glossary; Dobzhansky, 1962, pp. 147-151; 1970, pp. 101 ff; Ewens, 1969, chap. 2; Li, 1955, pp. 253 ff) that a tautology is involved in the theoretical relation expressed by the paired fitness/natural selection concepts with their single measure of residual relative frequency changes. Normally statements to the effect that “natural selection simply is differential reproduction” are made in contexts denying that natural selection is primarily a matter of “nature red in tooth and claw” (organisms most often compete in ways other than those involving direct aggression), or is primarily a matter of a “survival of the fittest” (differential reproduction rather than merely survival is what counts), or that it even always involves competition over scarce resources (intrinsic differences in reproductive rates can lead to shifts in the absence of scarcity, a point which Darwin himself downplayed). Such statements are not made in other words in contexts which imply a rejection of the adapted or fit concept. After all, the obvious design-like character of the major features of organisms was one of the major classes of facts Darwin’s theory was designed to explain. Heritable variations were not simply differentially perpetuated (at random for example) to Darwin, but differentially perpetuated

because of their different characters. “This preservation of favourable variations and the rejection of injurious variations I call Natural Selection” (emphasis added, Darwin, 1958, chap. 7). A similar historical justification for the inclusion of a concept of the competence of acts in learning theory, analogous to the adaptedness of organisms in evolutionary theory is not available however. Skinner’s way of handling this problem is therefore one possibility.

The alternative solution proposed here however not only will be shown to contribute to theoretical integration but is empirically justified. The evidence for such a solution is the evidence for what has been called the “relativity of reinforcement,” for the fact that what will serve as a reinforcer depends not only on the taxon involved, (which is to be expected from biological theory), but also on the act involved. (For reviews and discussions see Bitterman, 1975; Bolles, 1973; Hinde and Stevenson-Hinde, 1973; Houston, 1966, chap. 4; Seligman, 1972; and Shettleworth, 1972 (b).) For example, in rats shock avoidance will strongly reinforce running in wheels, weakly reinforce turning around, and not at all reinforce standing on their hind legs although the spontaneous emission rates of these behaviors in the situation are initially approximately equal. Sevenster (1968) has shown that the opportunity to court a female for sticklebacks is a relatively inefficient reinforcer of stick-biting as opposed to swimming through a hoop. Shettleworth (1973) has demonstrated that food for hamsters is a relatively inefficient reinforcer of face-washing, scent marking, and scratching with a hind leg as opposed to bar-pressing, scrabbling, digging, and open-rearing. If an event’s ability to reinforce depends on the response class, it suggests the need to relativize the reinforcement concept along the lines suggested here. In evolutionary biology it is considered legitimate to say that a positive selection pressure has the effect it does because of the “adaptedness” of the organism involved despite the

fact that what is adapted varies greatly. I am suggesting that evidence about the relativity of reinforcement makes it legitimate to say that a reinforcer has the effect it does because of the “competence” of the act involved despite the fact that what is competent varies greatly.

Reinforceability probably depends upon many particular neurophysiological qualities of the act involved including perhaps its “location” for example.

Other explanations for the relativity phenomenon which have been put forward tend to boil down to the same thing that has been suggested here. Shettleworth (1973), for example raised the possibility that in the attempt to reinforce face washing in hamsters with food, a tendency to approach the feeder might also be being classically conditioned in the situation, specifically to the initiation of face washing. Dunham (1977, pp. 119-122) has offered a similar “incompatible responses” interpretation of Moore’s and Sevenster’s research. But if one response is being instrumentally conditioned in a situation and another is being classically conditioned, such explanations imply that one response cannot “compete with” another in the situation, and that in turn implies that one is more competent than the other in the situation. It is important the competence not be confused with biological adaptedness concretely. As is readily observable particularly in human beings, not all learned behaviors are biologically adaptive. An act may be competent without being adaptive or adaptive without being competent. Avoidance of this confusion is particularly important because the most widely favored explanation for the relativity of reinforcement has been -- from the point of view of learning theory strictly speaking -- a non sequitur. It has been widely hypothesized that such relativity is due to evolution. Just as the efficacy of drive-reducing events or “broad-spectrum” reinforcers such as food, water, and opportunities for sexual activity - able to reinforce a large variety of response classes, is

ultimately attributable to evolution, it makes evolutionary sense that hamsters have evolved so that various locomotory responses are more reinforceable by food than are grooming responses for example. This sort of explanation is almost undoubtedly correct as long as one keeps in mind that adaptation is not perfect, i.e. evolution does not maximize adaptation but maximizes it only opportunistically. (Adaptation is constrained by the opportunities afforded by mutation and recombination, and the presence of many genes in a population represent a compromise resulting from the fact that their contribution to or detracting from the fitness of the whole organism varies with the rest of the genome.) But keeping in mind the non-competing distinctions in Biology among evolutionary, genetical, developmental, and physiological explanations for a trait, and given the fact that traditionally the psychology of learning has tended to concentrate on the latter, on a kind of molar physiology of behavior, then from the latter point of view, such an explanation is also a non sequitur. Pointing to the adaptive significance of a particular reinforcer or a particular case of selective reinforceability is to provide part (the evolutionary part) of an answer to the question of how such a state of affairs has come to be. It is not to answer the question of how it is that it is in physiological or psychological terms. Natural selection is not selecting acts one by one for perpetuation within a single life cycle, reinforcement is. But we must also realize then that a parallel distinction exists. If natural selection is not selecting acts one by one for perpetuation within a single life cycle but reinforcement is, then acts are not being selected within a single life cycle because of their adaptedness but because of their competence.

It is surprising that in previous papers on the learning/evolution analogy such an elementary point of comparison as the fact that both the effect and the natural selection principles have been accused of being tautological has not been explored. I will freely admit that the

“solution” to the problem of the basis of reinforcement presented here derived from such a comparison, could, with much justification, be seen as not much of a solution. At least it could be seen as only suggestive in a negative sort of way, as suggesting that no general behavioral solution exists and that instead the answer may lie in the (probably multiple) detailed neurophysiological properties of the act involved and the (probably multiple) detailed neurophysiological properties of the reinforcer involved in each case. But there is more to the proposed solution than that.

Just as in the study of evolution where the simple fact that organisms do vary in their adaptation to a setting is of greater theoretical interest than the diverse facts about what will serve as a positive selection pressure for this or that organism and why, in the study of learning, the simple fact that acts do vary in their competence in a setting may be of greater theoretical interest than the diverse facts about what will serve as a reinforcer for this or that act and why. The concept of competence used analogously to adaptation in evolutionary biology is designed to permit the now generally known fact of the relativity of reinforcement to be put to work in learning theory. The very existence of such a property bears important implications for other problems to be discussed and its utility therefore must be also judged in that light.

4. A single learning process

The problem of the relationship among different procedures known to yield learning is an old and difficult one in psychology (Seward, 1970) which has become more acute in recent years due to the discovery that autonomic acts can be instrumentally conditioned (Kimmel, 1974; Miller and Carla, 1967), and that skeletal acts can be classically conditioned as for example in

auto-shaping (Brown and Jenkins, 1968; Williams and Williams, 1969). Such discoveries have led to a renewed interest in the possibility of a theory which reduces the effects of certain procedures yielding learning to the processes believed to be involved in others, or at least in a single process theory in the sense that it is maintained that often, if not always, the different processes accompany and interact with each other. (For various viewpoints see Catania, 1979; Mackintosh, 1974; Rescorla and Holland, 1976; Rescorla and Solomon, 1967; Schoenfeld, 1966; 1972; Schwartz and Gamzu, 1977; Terrace, 1973). The bearing of the evolutionary analogy on this problem has been discussed (Staddon and Simmelhag, 1971; Staddon, 1975) in the context of superstitious or temporal classical conditioning procedures yielding terminal and interim activities. The application of a reinforcer at fixed intervals independent of either responding or any particular CS leads to the interval on both sides of the reinforcer coming to be littered with responses. As previously discussed Staddon and Simmelhag (1971) argued that phenomena analogous to superstitions are impossible in evolution but they also went on to propose that possibly reinforcement like natural selection is “purely selective” or “suppressive” and that much of the learning in such procedures may be analogous to “principles of variation” in evolution. Later Staddon (1975) expanded on this, suggesting that R-S contingencies may only restrict the range of variation in behavior possible while most learning phenomena may be attributable to practice and to stimulus including temporal stimulus contingencies.

It has already been pointed out that phenomena analogous to superstitions are quite conceivable in evolution. Moreover, the view that natural selection is relatively unimportant in evolution because it supposedly only eliminates while mutation and recombination create is one of the oldest fallacies about evolution (Simpson; 1967, pp. 202ff). One of the most important

reasons it is a fallacy is because past selection is a major determinant of the kinds and frequencies of subsequent mutations. A population with one mean height will give rise to a range and relative frequency distribution of mutations affecting height quite different from that of a population with a very different mean height for example. The creative role of natural selection in evolution has been defended on grounds precisely analogous to the creative role of shaping in learning in other words. In both processes, variation and selection contribute to the creation of new kinds of things. The second component in this use of the analogy, that perhaps most behavioral change is attributable to practice and to stimulus including temporal stimulus contingencies does not of course provide a process-type explanation of their effects.

Here then the evolutionary analogy will be applied to the problem of the relationship among procedures producing learning in a different way. In each of sections one and two it was shown how one inaccurately drawn or omitted aspect of the evolutionary analogy could clarify one problem in learning theory. This next step is somewhat more complicated. It requires that we simultaneously consider two additional aspects of the analogy (discriminative control as induction and the genotype/phenotype distinction) before they can be put together and put together with what has been previously said for their significance to be appreciated. The persistence required will be rewarded however because an accurate and complete evolutionary analogy provides what is ultimately a very simple theory of how all of the elementary procedures known to yield learning could be aspects of one single learning process.

4.1 The evolutionary analogue of discriminative control

While discussing the learning/evolution analogy, Gilbert (1972) remarked that there is no evolutionary analogue of the process of discriminative control in learning. This is in part at least false. In most respects the concept of environmental influences on an organism's phenotype, of environmental "induction" in evolution, is analogous to discriminative control of an act in learning. Natural selection works in the same "conditional" fashion as do reward and punishment. The phenotype of an organism is a joint product of its genotype and of environmental influences acting on it as it undergoes a development or ontogeny. Natural selection finds a phenotype adapted or fit and selects for it - hence the genes which gave rise to that phenotype become more probable in the next generation. Does that mean then that the phenotype becomes more probable? It does, but only if whatever "inductive" environmental influences which were in part responsible for giving rise to that phenotype in the last generation, are again present. What actually appears, the phenotype, is conditional upon external inducers in evolution and discriminative stimuli in learning in an analogous fashion. The only disanalogy is that in learning, we tend to view discriminative stimuli as initiating a learned act. Traditionally for example it was believed that reinforcement strengthened bonds between discriminative stimuli and observable behavior. The initiation implication remains in Skinner's theory, although the phrase employed, that discriminative stimuli "set the occasion" for an operant is ambiguously process-free. In evolution on the other hand, it is realized that an organism's genotype lies behind and antedates its observable features or phenotype. A genotype is not initiated environmentally, it simply appears as a consequence of a past history of selection. Subsequently, an organism undergoes an ontogenetic or development process and during this process inducers act, participating in determining the form the final structure takes or its phenotype. Staddon raised the

possibility of an act possessing an analogue of an organism's genotype and phenotype (1975, pp. 40, 54 and 90) but left this possibility relatively undeveloped. Let us imagine that it does and that discriminative stimuli play a role in the learning process analogous to that of environmental inducers or environmental influences on a phenotype in evolution.

Neurophysiologically such a situation could be realized in the following way. Acts could begin with the firing of one or more pacemaker neurons or coupled groups (sometimes called central pattern generating networks). Such uninitiated beginnings, manifesting an endogenous rhythm, would represent for an act, the analogical equivalent of an organism's genome. When more than one individual and/or group is responsible for initiation, they might fire or burst at the same or at different times relative to each other. (One may think of the multiple beginning situation as analogous to the multiple genes in an organism.) It is worth noting that such multiple starting points could provide "structure" for example syntactical or grammatical structure to an act which is neither a consequence of "chaining" nor of the occurrence of a temporally extended unit in which earlier stages cause later ones. The act would then subsequently "develop" as these initial firings activate other neurons resulting eventually in a particular pattern of muscular and glandular activity - in an act's "phenotype" or topography. The neurons involved in later stages of the development of the act could require "horizontal" environmental input as well as pacemaker/coupled group input for their activity and which pathways are followed then could vary with this horizontal input. (The same theory could be realized without a sequence of cells if the initiating neuron(s)/coupled group(s) possessed an endogenous rhythm but also required particular synaptic input at a specific stage or stages of that cycle to fire or burst). In either case, reinforcement would act by permitting such endogenous

cycling to continue and by increasing its frequency. But because of the existence of the development process, and the need during it for particular discriminative stimulus input, such reinforcement will also increase the probability of occurrence of a particular act topography or phenotype only if the proper discriminative or inducing environment surrounds the act's development on a subsequent occasion.

4.2 The evolutionary analogue of sensitization and habituation

Thus far then we have two kinds of environmental influences on an act in learning analogous to two kinds of environmental influences on an organism in evolution. The first kind, reinforcement and punishment, occur after act phenotypes, and act selectively, depending on the competence of the act vis-à-vis them. They are analogous to natural selection. The second kind, discriminative stimuli, occur before act phenotypes (but not before their genotypes), and act inductively, actually permitting or preventing act development and causally contributing to its eventual topographic characteristics. They are analogous to environmental inducers or environmental influences on an organism's phenotype in evolution.

The problem in integrating the basic known learning procedures into a single process theory has been the problem of what kind of a process could be going on when a stimulus which occurs prior to the observable features of an act, which appears to initiate an act, can, at the same time, be a reinforcer or punisher of that act. This occurs in sensitization and habituation or example, as well as in classical conditioning. If however we are dealing with a full evolutionary analogue, if such stimuli rather than initiating, eliciting, or setting the occasion for an act are inducing its development, all of these procedures can be seen as aspects of a single evolution-like

process. This is because in evolution an inducer can at the same time be a selection pressure.

Sensitization and habituation are perfectly comprehensible in the light of the now fully extended analogy. In evolution, selection pressures can be seen to fall into two groups. A selection pressure can simply be a thing or event in the presence of which one organism is more or less adapted than alternative kinds of organisms, such as the presence of soot in the classical selection of dark over light moths case. Such a selection pressure occurs after the phenotype or at least after the relevant aspect of the phenotype has developed. But a selection pressure can also be an inducer, a thing or event which inductively alters phenotypes if it induces development in an adapted or maladapted direction relative to all other influences present, such as normal maternal influences as opposed to the influence of thalidomide on the development of a placental mammal for example. And inducers of course occur prior to the phenotypes they induce.

In learning in some cases the simple occurrence of or the “elicitation” of an act makes it more probable (sensitization), and sometimes less probable (habituation). A commonplace instance of habituation occurs, for example, when a fairly loud noise induces a startle response in a dog. A series of such presentations is likely to result in that act being elicited with less and less reliability. It has never been easy to find a place for habituation and sensitization in traditional theories of learning couched in terms of reinforcers creating or strengthening stimulus-response connections, or in terms of stimulus control events as eliciting or setting the occasion for behaviors. In a learning theory which conceives of acts being perpetuated at different rates because of the competence-reinforcement nexus, and which conceives of stimulus control events as inducers however, the situation is different. If inducers affect the topography of acts, then those effects can be in the direction of making the act more or less competent in the general

circumstances present. Therefore an event of induction by itself can alter future probabilities of occurrence and we have sensitization and habituation.

The application to the problem of interim activities in superstitious procedures is straightforward. When a reinforcer is applied on a fixed time schedule independent of both behavior and other stimuli, there is no reason why such a reinforcer could not be both selectively “superstitiously” reinforcing an act which precedes it (a terminal activity) and “inductively” reinforcing (or sensitizing) an act which follows it (an interim activity) at the same time. For that matter it could be selectively and inductively reinforcing a number of different acts which have slightly different life cycle lengths and/or are currently at different stages of development and which therefore will emerge phenotypically at slightly different times during the terminal and interim period respectively.

4.3 The evolutionary analogue of classical conditioning and conditioned suppression

[Insert Figure 1. about here]

Figure 1. illustrates how the various learning procedures are related to each other according to this theory. An act is illustrated as a horizontal line with the genotype/phenotype distinction indicated. Reinforcement and punishment are illustrated as arrows with solid shafts, inducers as arrows with broken shafts. In Figure 1a instrumental and operant reinforcement and punishment are shown as operating selectively, post-phenotypically. In Figure 1b instrumental and operant discrimination training are illustrated. An inducer, or S^D has been added. In Figure

1c sensitization and habituation are illustrated. The arrows are shown as acting inductively i.e. as in the pre-final phenotype position, but with a both broken and solid shaft, indicating their simultaneous status as reinforcers or punishers. (They are analogous to the evolutionary cases of induction to an adapted structure by maternal influences on the development of a placental mammal, and hence selection for, or induction to a maladapted structure by thalidomide, and hence selection against.) What about classical conditioning and conditioned suppression? They are illustrated in Figure 1d. They are like instrumental and operant discrimination training (1b) in having a distinct inductive event, the CS, but like sensitization and habituation (1c) in that the reinforcer or punisher is acting inductively, pre the final phenotype. The situation is one in which a sequence of two inducers are operative, the second one of which is inducing act development in a competent or incompetent direction respectively. In other words, classical conditioning and conditioned suppression are seen as discriminative forms of sensitization and habituation.

4.4 Summary

The argument made here about the implications of the evolutionary analogy then is inordinately simple. Learning like evolution is apparently a selection process. But one cannot talk about selection without talking about the adaptedness of the things subject to it. It is the other side of the same coin. And this is especially true when we know empirically that the things subject to it are differentially adapted to different selection events. This is the significance of the relativity of reinforcement for what is going on here and now in the organism (although particular

cases of relatively almost certainly have additional significances in other context, for example in an evolutionary context). As long as it was believed that reinforcers were uniformly transituational this other side of the “selection” coin could be ignored. Now that we know that this is not the case, it cannot continue to be ignored, or to be dismissed as requiring only an evolutionary explanation (nobody would argue that because hearts evolved we do not have to explain why they tick), or to be allowed to sit as a widely acknowledged fact alongside learning theories without being absorbed into those theories in some way. The only departure from a straightforward extension of the analogy has been the suggestion that the term “competence” be substituted for adaptedness in order to avoid confusion with biological adaptedness concretely. Once one becomes aware of this other side of the selection coin, and once one adds to it the genotype/induction/development/phenotype concepts, it can serve as a lynchpin tying together different procedures yielding learning into a single theory of the process involved. These terms too it must be stressed are being used as analogies. Actual genetic influences on behavior are not being referred to. The current array of neurophysiological rhythms are presumably a consequence of evolution heredity development and past learning. Once analogues of these concepts are applied to acts and added to the analogue of adaptation, the elementary procedures yielding learning can be combined into a single learning process in the following way. In evolution change can be wrought by changing the selection pressures, the environment. But it can also potentially be wrought by changing the organisms. If I change the structure of an organism with thalidomide for example, because such phenotypic changes have consequences for survival and reproduction, changes in future relative genotype frequencies are brought about. Hence if all stimuli which occur prior to the observable features of acts and which influence the

future course of events are acting inductively, sensitization, habituation, classical conditioning and conditioned suppression can be seen as procedural variants of the same selection process involved in reinforcement and punishment.

Note also that when we say an act can be made more adapted or competent inductively, its adaptedness to or competence with respect to any single specific reinforcer identifiable to us is not being referred to. To be useful the concept of the adaptedness or competence of act has to become just as real to us as is the reinforcing effect of stimuli. As it is now we use the term reinforcer in this way. When a stimulus is applied, and an act preceding it as a consequence becomes more probable, we say the stimulus acted as a reinforcer, most often in total ignorance of just what the properties (e.g. topographical or neurophysiological) were of the act which made it adapted to that reinforcer. Similarly, when we apply a stimulus, and an act following it as a consequence becomes more probable, we have to be free to say that we inductively made the act more adapted or competent even though we are, for now at least, ignorant of what the properties were of the environment that made it a positive selection pressure for the act thus altered. In summary then

- with reinforcement and punishment we select for and against acts
- with sensitization and habituation we make acts more or less adapted or competent
- with discriminated reinforcement and punishment we select for or against acts influenced by inductive stimuli which will therefore be more or less probable in the future as long as the same or similar inducers are present to influence their early development again
- with classical conditioning and conditioned suppression we make acts influenced by inductive stimuli more or less adapted or competent which will therefore be more or less

probable in the future as long as the same or similar inducers are present to influence their early development again.

4.5 Some evidence and implications

This way of seeing the relationship among different learning procedures modelled on the evolutionary process gives rise to many questions concerning evidence and how it relates to other theories. Since there is not space enough to explore these in detail here, a list of questions and mostly brief tentative answers are presented in this subsection. The matching law and avoidance are discussed briefly in section 5. Anyone less interested in the subtleties of how the theory presented bears on the relationship among different learning procedures, which is still our theme, might wish to proceed directly to section 6. where the final fundamental point is made, namely, that the analogy suggests that learning may be both a selective process, attributable to rewards and punishment, and something of a cognitive one as well.

4.5.1 Is there evidence for the existence of act “genotypes” of the type hypothesized?

Not a great deal of direct evidence, no. On the other hand, the existence of endogenously rhythmic behavior is well established. So too it would appear is the existence of spontaneously oscillating pacemaker neurons and couples groups (Grillner, 1975; Hoyle, 1973; Huber, 1975; Strumwasser, 1974; Wilson, 1970). It has also been discovered that endogenous firing and bursting rates can be altered by synaptic input (Parnas, Armstrong and Strumwasser, 1974; Parnas and Strumwasser, 1974). The study of such phenomena has proceeded largely in isolation from the study of learning however. The suggestion made that reinforcement and punishment are based on the modulation of such events is similar to Pringle’s (1951) paper.

4.5.2 Is there evidence that acts “develop”?

In a common sense way, once said this is obvious. But there is a deeper kind of hint in the analogy between the biogenetic law and the gradient of reinforcement. As evolution proceeds, more change tends to take place in the later stages of individual ontogenies than takes place in earlier stages. The biogenetic law (not recapitulationism!) stating that related groups of organisms tend to be more similar earlier in development than they are later is an approximately true generalization although exceptions certainly exist. It tends to hold because development is a complex process of interaction among cells and parts of an embryo. What happens later in the process therefore is in part, conditioned, limited, or determined by what has occurred earlier in the process. A mutation which has substantial effects on an embryo early in development then would be, on the average, less likely to be adaptive than one whose major effects appear later because of the more complex ramifications of the former. The most successful way to change the outcome of a complex process like development is to redirect its course as it proceeds rather than to interfere in major ways near its beginning. Waddington for example has expressed this view:

it is only to be expected that evolutionary alternations are much more likely to affect the later stages of development when comparatively minor features are being formed, and to leave intact the earlier steps on which all the later stages must depend. (1956, p. 60)

The gradient of reinforcement or of temporal discrimination (Dews, 1962; Hull, 1952; Zeiler, 1977, pp. 213-221) would seem to be an analogous phenomenon. When a lengthy complex act is being learned, say a maze with a series of choice points, we tend to find that improvement occurs or mistakes are eliminated in a backward direction. It takes fewer trials to learn the correct choices later in the maze than it does to learn those closest to the beginning of the maze.

A similar phenomenon is displayed in frequency terms, both in the acceleration in responding observed within intervals on fixed interval schedules, and (patterned differently), within the series of responses bounded by successive reinforcers on fixed ratio schedules. Note that in evolution this “more change later than earlier” applies within (not between) entities being selected for as units. The analogy suggests the possibility of using such a theoretical expectation to help identify what entities (like numbers, rates, and/or temporal patterns of similar or different responses) are actually being selected for as units on schedules which do or do not explicitly programme them. Herbert Simon (1962) once concluded that any evolutionary process should show an equivalent of the biogenetic law. The fact that the learning process seems to do so suggests that, as in evolution, it is “developing” entities which are “evolving” in learning.

4.5.3 Is there evidence that controlling stimuli act “inductively”?

The theory of controlling stimulus events as inducers of act phenotypes in all learning procedures rather than as initiating act genotypes implies that if an act (phenotype) can be brought under the control of a particular stimulus, it must, on its very first appearance, have been induced by that stimulus. The mere presence of an S^D or a CS just prior to or during a reinforced act is not enough for control to be acquired in other words, it must already be an inducer of the relevant phenotype. It is just that before training the relevant genotype was infrequent while after it has become more frequent. This means that the theory predicts that not all stimuli should be “attachable” to all acts. The relativity of reinforcement should extend to stimuli behavior is to be brought under the control of. Kriston, Lindauer and Menzel have shown that bees can most easily use olfactory clues, less easily use colour cues, and least easily use pattern cues when learning to get to and from places (cited in Hinde and Stevenson-Hinde, 1973, p. 11).

Shettleworth (1972a) has shown that chicks more readily use visual than auditory cues when learning not to drink or peck. Garcia and Koelling (1966) have shown that rats can learn to associate tastes but not lights or sounds with a digestive upset and the reverse with electric shock. Dobrzecka, Szwejkowska and Konorski (1966) have shown that dogs can use the quality of tones but not their spatial locations when learning go and no go discriminations. Just as the basic relativity of reinforcement is anomalous in terms of traditional theories of the learning process, the fact that within an act/appropriate reinforcer class, this relativity extends to stimuli the behavior is to be brought under the control of its anomalous. With a three cornered interpretation of the act however, involving a genotype, induction and a phenotype, it is clear that there need not necessarily exist an act genotype requiring stimulus B for its development into phenotype P to be made more probable, just because there exists an act genotype requiring stimulus A for its development into phenotype P.

Now of course, just as with the basic relativity of reinforcement one could appeal to more antecedent explanations, to evolved and inherited tendencies to associate particular stimuli with particular responses for example. But again, as with the basic relativity of reinforcement, whether or not something has biologically evolved and is inherited, there must also exist some current psychological/neurophysiological reason why it is so. One could of course appeal to concepts like “attention,” but since in common sense usage attention is something that organisms do, it would at least have to be altered to something that acts do because apparently different acts “attend” preferentially to different stimuli.

An additional category of evidence for the theory of controlling stimuli as inducing would be provided by showing that as the stimulus conditions under which an act was learned are

altered, not only is response probability altered, but the phenotypic properties of the acts which occur, such as amplitude for example, are distorted. In evolution it is the fact that alterations in the inductive environment distort phenotypes, that they are teratogenic, which accounts for their influence on the probability of development being completed. In stimulus generalization tests both probability and amplitude typically declined. But the theory of controlling stimuli as inducing would predict that properties like response amplitude would be distorted, not necessarily that they would always decline. The details of what occurs would depend largely on the degree of the relative contribution of genotype and environment to the phenotype, and the kind of contribution of environment to phenotype, i.e. whether it was making the phenotype more intense or less intense. Phenomena suggesting such idiosyncratic effects of altering inductive circumstances have been observed as discussed as sensory superstitions and the effects of an eccentric stimulus (Staddon, 1975, pp. 73-75), and are part of the phenomenon of behavioral contrast (Schwartz and Gamzu, 1977). All of these however have been studied using response rate as the dependent variable, and because of the complexity involved in interpreting response rate (or reproductive rate) as a phenotypic property of a response, we shall not go into these examples further.

4.5.4 The operant/respondent distinction

It follows from the single process theory presented that an absolute emitted operant and elicited respondent distinction is rejected. Such a rejection is in fact already widespread. Rationalist (or endogenous) and empiricist (or environmental) aspects of behavior can be encompassed by merging them at the level of the execution of each single act instead of dividing

the world of acts into two classes. Each act may contain an endogenous (“genotypic”) and environmental (“inductive”) component. Respondents and operants then can be seen to stand at the opposite ends of two continua. A respondent such as the human patellar reflex is seen as an act which is “available” extremely frequently because its endogenous rhythm is very rapid, but as one which also requires a highly specific set of inductive circumstances for its development. Its position at an extreme end of these two continua is what would give rise to the appearance of pure environmental elicitation. As Catania (1979) has put it it is highly probable in the presence of a particular stimulus and highly improbable in its absence. An operant such as the rat’s bar press is seen as an act which is much less frequently available because its endogenous rhythm is slow, but as one which also is broadly tolerant of a diversity of inductive circumstances. Its position at the opposite extreme of these two continua is what would give rise to its voluntarily emitted appearance. For this reason I have preferred throughout to speak of acts rather than respondents and operants or even responses.

4.5.5 Contingencies

It has been common to describe the difference among different learning procedures in terms of the contingencies involved - temporal, SS, RS, and SRS contingencies. But in a selection process there is no such thing as complete contingency or dependence. Even in simple operant arrangements with automatic equipment set to always reinforce a bar press and never reinforce in its absence, the occurrence of the reinforcing stimulus is not solely contingent, dependent, or consequential on the behavior. It is also dependent on or partially caused by the fact that the experimenter set the equipment in a certain way, on the fact that it continues to work

in good order and so on. The point may seem trivial but is important in understanding selection processes. In evolution for example, selection is never wholly dependent on the structure of the organism or wholly dependent on the environment. Both exert an independent influence - peppered moths and soot for example were both required for the classic industrial melanism case. In instrumental or operant procedures contingency has been shown to be neither necessary or sufficient for learning to take place. The acquisition of superstitions shows it is not necessary. One could of course maintain the contingency or consequential characterization by classifying the results obtained from fixed time schedules with all other effects of stimulus presentation, i.e. by implying that terminal activities are classically conditioned. But there is no compelling reason to believe that the process involved in the learning of terminal activities is any different from that involved in the learning of a normally “contingently” reinforced behavior. Contingency is not necessary then, nor is it sufficient. It is not sufficient because some degree of temporal contiguity is required, contingency with delays of hours would normally not do for example. It is also not sufficient because it omits any consideration of the relativity of reinforcement, a contingent stimulus may reinforce one behavior but not another. R-S contingency then does not appear to be a particularly useful classificatory principle.

The evolutionary analogy would suggest that in addition to the requirement that the act be “adapted” to the reinforcer topographically (which could involve multiple neurophysiological properties of both act and reinforcer), the key ingredient is temporal contiguity combined with the reinforcer’s occurrence at the right “life cycle stage” of the act. Birds for example need nest building material in the spring, not at any old time at all. It is this “right life cycle stage” or phasing requirement that accounts for the fact that an act established initially on the same VI

schedule in two cases, will be maintained better subsequently on a VI schedule than it will be on a VT schedule (Catania, 1979, pp. 177-180). In the latter case, the greater variability in the distribution of time intervals separating act and reinforcer means a lack of consistency in the phasing of the reinforcer relative to the life cycle of the act, and the procedure is less efficient, more difficult for the act to adapt to, as a consequence. But this difference does not “distinguish” the difference between accidental correlation and causality as Catania (1979, p. 180) maintains. An event potentially can be both temporally contiguous and occur with some consistency at the right “life cycle stage” if it proceeds from an independent rhythmic source, with identical frequency and phasing. Many endogenously rhythmic behavioral and other biological processes are now known to occur independently of the rhythms, often geophysical, with which they coincide (although they are often entrained to the latter). The high degree of response contingency obtained with automatic equipment than is effective because it is the simplest way of maintaining consistent temporal contiguity and phasing of the reinforcer relative to the act (Zeiler, 1977, pp. 204-5).

For purposes of theoretically understanding the process involved therefore, I believe that Schoenfeld has repeatedly put his finger on the most important difference between classical and instrumental or operant procedures, the temporal order of events. In classical procedures, the reinforcer initially precedes the act of interest while in instrumental and operate procedures the reinforcer initially follows the act of interest (Schoenfeld, 1966, p. 222; 1972, p. 53; 1973, p. 153). The problem then is how can a stimulus reinforce an act which it precedes? One suggestion made by Terrace (cited in Hilgard and Bower, 1975, p. 211) is simply to say that an US both elicits and reinforces. But this not only does violence to our notion of reinforcers and

punishers as selecting among what then must be antecedent behaviors, it also happens to be empirically false about half the time. When we apply a stimulus which elicits an act, sometimes that application does make the act more probable (sensitization) but in other cases it makes the act less probable (habituation). The theoretical key it seems to me is not to see such stimuli as eliciting at all, but as inducing. If inducing, if permitting or preventing act development by actually participating in determining the latter's topographic features, then such a stimulus can be inducing development in either an adapted (competent) or maladapted (incompetent) direction relative to all other influences present, and hence have an indirect selective effect.

4.5.6 Connectionism and stimulus substitution

It is clear that this is not a connectionist theory. Evolution does not create or strengthen bonds between inducers and phenotypes. Although it could, viewed in a particular way, look like that, this is a by-product of the fact that it alters arrays of extant genotypes. Nor is the theory then a stimulus substitution theory of the classical procedures. In evolution, pairing inducers sequentially will not lead, after a few generations, to the first acquiring, through "association," an ability to induce the phenotype normally induced by the second. However, we can make an act induced by the first more probable, by reinforcement. In the classical case it is just that the reinforcer is also an inducer, inducing act development in a competent direction.

Do not classical procedure often control for sensitization? The normal control for sensitization means seeing if the same result can be produced by applications of the reinforcer without its being correlated with the to-be-conditioned stimulus. But this correlation is necessary. We want to "inductively reinforce" an act induced by the to-be-conditioned stimulus.

It is not being claimed that classical procedures work solely via sensitization and habituation but that they are discriminated forms of these.

It is often asked why the two acts in classical conditioning, the one initially inducible by the reinforcer alone and the one inducible by the combined pair, are often different. The one and two inducer situation obviously can produce different phenotypes, but because of the common and important second inducer, they often overlap to a greater or lesser degree. By the time we arrive at test trials, or by the time the act has (by becoming more probable) come to fully antedate the reinforcer, the original act and the new one will of necessity be at least slightly different. This is because, even if they were relatively similar during training, the inductive environment of the latter stage of development is now altered by the non-presence of the reinforcer and the phenotype is changed at least somewhat by this. As well, of course, the act is now no longer being reinforced and will not continue forever to be induced by the conditioned stimulus without “repairing.”

4.5.7 Compared with other attempts at reduction

The most unique feature of the theory proposed is that it neither attempts to reduce instrumental or operant procedures to a classical process nor vice versa. Instead it reduces both procedures as well as sensitization and habituation to a new single process theory. In that sense it is a realization of the kind of solution hypothesized to exist by Schoenfeld (1972, p. 53). Other attempts at reduction or partial reduction typically seek complex ways around the procedural order of events traditionally believed to be operative at the beginning of training - in classical procedures, reinforcer then act, in instrumental and operant procedures, act then reinforcer (in

either case with or without an earlier signal). They do this by postulating that reinforcers in classical procedures are really working by selection or “backward” to employ a useful image, or that reinforcers in instrumental or operant procedures are really working via elicitation or “forwards.” In both types of reduction this is held to be the case because the reinforcer is really working on a different act, or a different or persisting instance of the same act. Some examples are Smith (1954), Rescorla and Holland (1976) and Terrace (1973). The complexity of these theories should be contrasted with the simplicity of the proposed single process theory. Nor is the addition of genotypes and the hypothesis that all controlling stimuli work by induction the same kind of attempt at reduction as these. Here reinforcers are located and are operative on both sides of observable acts (i.e. of act phenotypes) classical ones before, and instrumental or operant ones after. Both are held to work “in their own direction” so to speak, as traditionally thought, but via an interrelated process mechanism, the competence/reinforcement nexus. The reduction then is perfectly isomorphic. Cast as it is in terms of the evolutionary analogy it may appear more similar to our traditional understanding of instrumental or operant conditioning especially since stimulus substitution as a theory of the classical procedure is abandoned. But the reduction is isomorphic because traditional theories of the instrumental or operant procedure are abandoned as well, for example connectionism or theories based on an absolute operant/respondent distinction however defined.

5. Act classes, the matching law and avoidance

An interesting attempt to apply an aspect of the evolutionary analogy is Catania’s recent suggestion that behaviors be classified on the basis of their relatedness to each other via

“descent” in learning (1979, p. 359). The suggestion is made in the context of the contingency or consequential classification, so he begins with a group of behaviors with common environmental effects and then goes on to include their “offspring,” but only to the extent that they are similar to their “ancestors.”

The essential feature of an operant is the correspondence between a class or responses defined by its consequences and the spectrum of responses generated by these consequences. (1979, p. 121)

Now as a matter of fact there is no universal consensus in Biology on the ideal basis of classification - instead, three broad positions can be identified, the numerical, the evolutionary, and the cladistic (for an elementary review see Mayr, 1976, Pt. V). As a practical matter, most existing classifications in Biology are based on structural similarities and differences, modified at times by historical evidence about degrees of relatedness. As a practical matter too, most classifications of behaviors in the study of learning are differently based. They are based on having a similar effect on the world not at all on “topography.” This means that a typical instrumental or operant class encompasses three kinds of “evolutionary” phenomena -

- a) homologies or entities which are structurally similar because of common descent (for example, slightly varying nose bar presses)
- b) analogies or entities which are structurally similar because of common selection pressures (for example left turn and right turn derived bar presses which evolved into bar presses under the pressure of reinforcement through becoming left to right and right to left swipes at the bar), and
- c) ecological equivalents or entities which occupy a similar niche in the absence of great structural similarity such as Darwin’s finches and woodpeckers - several species of the former employed spines as tools to occupy a woodpecker-like niche. (A behavioral example would be

nose and limb bar presses).

The typical instrumental or operant behavior class generated in a laboratory experiment then is a very different kind of class than is a biological taxon based on any of the three broad philosophies of classification in Biology. None of this presents any great problem however, reinforcement, as does natural selection, simply accepts different solutions and differently derived solutions to the same problem.

A second aspect of the whole question of units in evolutionary theory does bear directly on an outstanding problem in learning theory however. Herrnstein's matching law stating basically that the relative frequency of each kind of behavior in a choice situation equals or matches the relative frequency of their reinforcements (de Villiers, 1977; de Villiers and Herrnstein, 1976; Herrnstein, 1961; 1970; 1971; 1974) is directly analogous to the most fundamental law of evolution - the Hardy-Weinberg equilibrium principle of population genetics. The principle states that the relative frequency of alternatives in a population remains constant unless something (particularly but not exclusively selection) acts to change it. The principle is treated as a convention, as an analytical truth. It is an inertial principle for gene pools used as a standard against which to measure events such as selection (Crow and Kimura, 1970, chap. 1; Dobzhansky, 1962, pp. 147-151; 1970, pp. 101 ff.; Ewens, 1969, chap. 2; Li, 1955, pp. 253 ff.; Mettler and Gregg, 1969, pp. 92-93). While the possibility has been raised that the matching law could be viewed in a similar way (Rachlin, 1971), this suggestion has been controversial and most frequently matching has been treated as an independently testable empirical generalization.

Herrnstein's formulation of the law of effect in relative frequency terms has relevance for the explanation of avoidance. At times organisms can learn a behavior by being punished, for

example with electric shock, for not performing it. The problem has been what then is reinforcing such an avoidance response. The conclusion that the non-occurrence of shock is the reinforcer (Fantino, 1973, p. 272) is not a very congenial one. As Schoenfeld has remarked:

The paradoxical point is sharpened if the reader asks himself how many stimuli are not occurring to him at this moment - would he care to say that his behaviour is limitlessly avoidance? (1972, p. 56)

However, if reinforcement and punishment effect relative frequencies of alternative acts in a population rather than the absolute frequency of a particular act, then in addition to the direct reinforcement and punishment processes we have been discussing to date, a set of indirect processes should exist. It should be possible to increase the probability of an act by punishing its alternatives (indirect reinforcement) and to decrease the probability of an act by reinforcing its alternatives (indirect punishment). Avoidance responses then may be learned via indirect reinforcement, they may be learned not because they are directly reinforced by anything, but because virtually everything else the animal does in the situation is punished.

The evolutionary analogy in other words suggests a competing response theory of avoidance. In the framework of various traditional learning theories, without an elaboration of the evolutionary analogy or without a matching law, concepts of competing responses were very much ad hoc additions without a systematic theoretical basis. Also, traditional competing response theories of avoidance have normally been two process theories - appealing to both classical (stimulus substitution) and instrumental components (e.g. Mowrer, 1947; 1960) which is not the case here. It is also true however, that within this framework, the learning of any particular avoidance response could be a two process event - the two processes being induction and effect rather than stimulus substitution and effect. For example, it would be possible for

shock to result in the learning of an avoidance response because it is punishing, in the effect sense, alternatives to that act, or, because it is inducing the development of alternatives in an incompetent direction.

Such an “indirect learning” or competing response approach can explain the most obvious and puzzling feature of avoidance learning, namely, the difficulty experimenters have experienced in training any arbitrarily chosen act with avoidance procedures. The theory predicts that the learning of acts via either indirect means should show a greater relativity or specificity in the response/reinforcer-punisher relationships than does learning via either direct means. This is because while one condition only is required for direct reinforcement or punishment (the response must be competent or incompetent respectively vis-à-vis the reinforcer or punisher), two conditions are required for indirect learning. In indirect reinforcement (avoidance) the punished response(s) must be incompetent vis-à-vis shock, but, in addition, the experimenter-specified avoidance response must exist as a member, as one of the alternatives extant in a specific act population or populations, namely, in the act population or populations being punished. If the avoidance response does not happen to exist as an alternative in this specific population or populations then shock applied indirectly will be ineffective in increasing the probability of this particular act. An analogous second condition would apply to the indirect punishment case. It is not surprising then, that avoidance learning turns out to be a very tricky proposition, and that it has contributed more than its share to the evidence establishing the relativity phenomenon. Bolles (1972) particularly has emphasized the “it depends upon the response” view of avoidance but has suggested another theory - that a particular response will be learned as an avoidance response only to the extent that that response is a species-specific

defense reaction” of the organism involved, such as freezing, fleeing, or hiding are for rats. It is probably no doubt true, as Bolles has suggested, that responses readily learned as avoidance responses tend to be those utilized in defensive ways in natural settings by the species involved. It would suggest, however, that the current psychological or physiological reason why certain specific responses are readily learned in avoidance situations while others are not is because the former exist as one of the alternatives available, as members of many different act populations. It is because of their broad “geographical-ecological” distribution, because of their presence in and ability to “compete” in many different populations, that they continually “pop up” as avoidance responses. It is not acceptable to attempt to substitute evolutionary for psychological/physiological explanations. Nor should the importance of the traditional psychological question of what it is that reinforces an avoidance response be down-played.

The matching law is not currently seen as able to explain this particular feature of avoidance learning. This is because it is held that matching holds among any set of behaviors arbitrarily specified by the experimenter. It is possible however, that our subject matter rather than solely some of our procedures is organized in populations of mutually exclusive alternatives. In that case indirect reinforcement i.e. the punishment of alternatives could not always be expected to work for an act chosen but not a member of the population involved.

Keeping in mind the broad similarity of the approach expressed in the Hardy-Weinberg equilibrium principle and the matching law, it seems obvious that an evolutionist would be astonished at the multiplying distinctions in psychology - reward and punishment, positive and negative, by instrumental or classical procedures, now direct and indirect. From his point of view they all amount to changes in relative frequencies and that is what matters. If we are

willing to assume as he does that our subject matter is organized into populations of mutually exclusive alternatives, that rewards and punishments affect relative frequencies, and, of course, that not all alternatives are present in every specific population, then it is possible to both understand how avoidance learning is possible, and particularly, how unusually difficult it has turned out to be to teach any arbitrarily chosen act with the kind of “indirect” procedures used in avoidance arrangements.

6. A cognitive reinforcement theory

One of the most exciting benefits of thinking through the evolution/learning analogy is that the evolutionary case suggests how behavioral and cognitive learning theory can be reconciled - positions which have traditionally been seen as incompatible. Moreover such a cognitive reinforcement theory provides a better explanation for phenomena such as place learning than do either traditional positions. The key to the reconciliation is the following combination. As in evolution, learning is essentially a process of selection, by reinforcement and punishment of blind variants. Unlike the traditional behavioral position however, this selection does not discriminate among actual behavioral topographies (act phenotypes) but among the act “genotypes” or programmes which give rise to them. Because environmental inducers (S^D s or CSs) also contribute to determining the final phenotype of the act, this means that a cognitive language about what is learned (for example an expectation), and about inducers (for example cues and warnings), is appropriate. Evolutionists too employ a cognitive language, organic genomes represent information or programmes they say.

The tricky part is that in both realms, such programmes often include alternative sets of

instructions for how to achieve adaptation or competence under alternative sets of circumstances.

This phenomenon is known in Biology as “regulative development” and can be demonstrated in learning by place learning experiments. This phenomena makes an even more cognitive language about what is learned, a purpose, seem appropriate. I hope to show below how much caution must be exercised when using this latter term however, and how cognitive theory as it stands, in non-selectionist form, is an unsatisfactory explanation of learning which is in fact an analogue of Lamarckianism. For convenience we shall discuss Lamarckian-style challenges to selectionist views of both processes in the historical context in which they arose.

6.1 Blind variation and selection

An important element in the pre-Darwinian Lamarckian theory of organic evolution was the belief that new hereditary variations occur in response to the demands of circumstances, that they appear in response to “needs.” In the post-Darwinian period it was found that new hereditary variations (or mutations) occur in all directions, that they are not oriented with respect to adaptaion, in fact, the majority of them tend to be deleterious. For a time among neo-Lamarckians, this issue became entangled with another, whether new hereditary variations are small or large changes. The latter position is often called “saltationism” or the belief that evolutionary novelty appears in large “leaps.” The modern consensus is that most evolutionary novelty appears in the form of small incremental change with some exceptions (the exceptions often involve structural changes in whole chromosomes or chromosome sets). For the most part, the question simply lost its importance as it gradually became clear that it really was irrelevant to the former and more important one. A saltatory change or leap is not necessarily adaptive and

vice versa.

An analogous set of issues rose in learning theory when the gestalt psychologist, Köhler, challenged Thorndike's reinforcement theory of learning and argued instead that learning is a consequence of "insight" (1925). During his famous studies on Tenerife during World War I, Köhler hung bananas out of the reach of chimps. After observing a period of what I now would call incompetent behaviour, i.e. behaviour ineffective in leading to reinforcement, one chimp suddenly climbed on a box and jumped from there, obtaining the fruit. Köhler emphasized the suddenness and the scale of the change in the chimp's behaviour and attributed his solution to insight. Here we have a problem in the realm of behaviour change or learning analogous to an element in the Lamarckian theory of evolution. The analogy suggests that the question of the "size of the leap" or the saltatory quality of the new behavioral variant is irrelevant to the fundamental problem. Of course new "competent" behavioural variants appear. Learning would never be observed if they did not. They may even on occasion be saltatory, or represent a major change with respect to what preceded them. The fundamental problem however, is whether or not new behavioural variants on a statistical basis are biased in the direction required for reinforcement to occur, are biased in the direction of competence. That the answer to this question is a negative one, analogous to that of evolutionary biology, is confirmed almost daily at the beginning of experiments involving choice. When two bars are presented to a group of animals, reinforcement requiring responding on one of them rather than on the other for example, we observe no statistical tendency for the correct bar to dominate initial choices. On a statistical basis, organisms are no more likely, when confronted with a novel problem, to produce new behavioral variants oriented in the direction of competence than to produce those not so oriented.

Indeed, it is difficult to imagine there being any other answer.

Response generalization can illustrate an important property of variation. Mutations in evolution are not random in any statistical sense that all possibilities are equiprobable. If they were, then mice might mutate into men as frequently as they mutate into slightly varying mice! It is well known in Drosophila for example that certain mutations occur at characteristic rates very different from those of other mutation. Response generalization can illustrate this important property of variation. Consider some quantitative property of a response, its amplitude for example. If we reinforce within some range of that quantitative property, the result is an increase in the probability of behavior within that range (not an equal increase for all values, even for those reinforced an identical number of times, because of the relativity of reinforcement with respect to varying values of a quantitative property), and also an increase, although a lesser one, in the probability of behavior outside the range. This latter increase tends to decline as one moves further away from the reinforced range on either side. Such a hypothetical gradient of response generalization is illustrated for example in Catania (1979, p. 124). Response generalization is typically attributed to a “spread of effect” from reinforced responses to similar, but non-reinforced responses. Now evolutionary theory would not attribute such a phenomenon to the spread of the effects of selection to organisms beyond those actually selected for! Instead, it would attribute such phenomena to mutation. A parental population gives rise to an offspring population which is not identical in all respects. Such gradients are often distorted in various ways. They may decline raggedly or more rapidly in one direction than the other for example. No simple response generalization or spread of effect explanation bears directly on such

distortions. But they illustrate the important point made above about mutations - that they are not random in any statistical sense, only “blind” as Campbell (1974) has put it. They do not tend, as a statistical body, to be oriented in the direction of adaptation, or here of competence. Going back to the two bar choice situation again, “preferences” may certainly be displayed for one bar over another in individual cases. All that evolutionary theory, and a learning theory which is an analogue of it requires, is that such preferences show no overall tendency to be correct. Selection rather than mutation in evolution, and reinforcement and punishment rather than variation in learning, are what orient change in the direction of adaptation and competence respectively. This is the fundamental anti-Lamarckian point, and questions about occasional large scale (saltationist or insightful) variation is really irrelevant to it. Failure as well as successes occasionally occur on a grand scale!

6.2 Place learning and regulative development

The same Darwinian/Lamarckian problem arises again in a more complex form in the cognitive or “what is learned” problem. The distinction has been variously termed movements versus action, movements versus achievements, molecular versus molar behavior, muscle-consistent versus object-consistent responses (for a review see Kitchener, 1977). This was raised as a problem in learning theory particularly by Tolman and demonstrated in place learning experiments. Bindra (1978) has recently revived this issue arguing that “motor-equivalence” is the single most important obstacle to a S-R reinforcement theory of learning and requires its replacement by a “perceptual (incentive)-motivational framework.”

An important objective feature of acts which Tolman experimentally demonstrated was

that they behave within limits in an equifinal fashion (our term not Tolman's). Within limits, we can interfere with their development by altering the conditions under which they occur, and they often produce the same result (reinforcement) anyway. Consider the following experiment. We place a rat in the south start box of a four arm maze arranged in the form of cross with food at the end of the east arm. After a few trials from south, the animal learns to turn right immediately at the choice point. Now place the rat in the north box and see what it does. Related experiments (Tolman, 1946; for a review see Restle, 1965) have shown that chances are that a rat, after such training, when placed in the north box will immediately turn left at the choice point thus obtaining the food. The argument that one might be tempted to make in a simple form would state that reinforcement cannot have been responsible for learning here because by definition reinforcement is an event which occurs after a response and increases its future probability of occurrence. But we reinforced turning right, not turning left - we reinforced turning right but the rat has also learned to turn left apparently. (This type of experiment exists in many variations. Instead of requiring the act to reorient its structure during execution from straight then right to straight then left, we could for example have left the food location unaltered but required different motions from the beginning, swimming instead of running say.)

Some reinforcement theorists have tended to dismiss this type of experiment by saying that there is nothing in reinforcement theory which says that rats will learn responses rather than places. I think however one has to admit that at least in its simplest, classical, stimulus-response connectionist form, with reinforcement seen as creating or strengthening "bonds" between stimuli and responses, that it would indeed predict this. If certain muscular movements involved in turning right were that occurred and were reinforced, then that and only that would be what the

animal should learn according to the traditional view.

Some have argued that Skinner's (1935) mode of defining operants in terms of their effect on the environment rather than their morphology or topography (a bar press is a bar press whether performed with nose, limb, or posterior for example) solves the problem (Epstein, 1978, p. 57). It is true of course that in a normal instrumental or operant experiment, the class of behaviors reinforced procedurally, includes all those performed which have a particular consequence. It is false however to imply that by such reinforcement, the class of all behaviors which could have a similar consequence is thereby increased in probability. If after learning, the only path to a goal is cemented up, it will obviously not be traversed or circumvented for example. Nor do we have to go to such extremes. It can readily be shown in such experiments (Restle, 1965) that if we alter the inductive influences or cues present sufficiently, the act will fail to adapt, and reinforcement will not be achieved. It is patently false therefore to say baldly and without qualification that organisms learn to "achieve consequences."

Restle (1965) has offered an analysis of these experiments which many consider to be definite (Catania, 1979). Place versus response learning is simply a matter of discrimination learning, of which cues have acquired control over behavior by consistent association, and which and what proportions of them are omitted, altered, or added in the testing situation. According to this analysis in the experiment described the testing would be pitting intra and extra maze cues (place learning) against muscular feedback cues. In other situations it is different. For example in cases where the maze is rotated through 180 degrees instead of the rat being moved to the north box, we are pitting extra maze cues against feedback and intra maze cues. If certain cues were not available they could not be used, for example, if the maze were covered from the

beginning extra maze cues could not be used.

Now clearly what discriminative stimuli are available and conditioned is part of the question, but the above analysis essentially begs the question. In the simple experiment described for example, if place learning resulted, Restle's analysis would simply say that either or both of internal and external maze cues had acquired control. But that control would have to be of the sort commanding "approach me" and the fundamental issue is how this sort of stimulus control is possible. If stimulus-response connections or actual behavior is what is acquired via reinforcement, how can stimuli acquire control of this locational sort which requires different motor responses in the training (turning right) and in the testing (turning left) situation. Similarly in the maze rotation experiment. If internal maze and feedback cues have acquired control, the question does not arise since a right turn is required in both the training and testing situation. But if the results show that external maze cues have, then the same problem exists. If the animal approaches the same location in the room during testing as during training, then these extra-maze stimuli are apparently able to command turning right in the training situation and turning left in the testing situation. While reinforcement theorists continue to largely ignore this basic problem, a whole generation of cognitive psychologists are coming to reject reinforcement theory largely because of this "action" problem.

Instead of these versions, I believe that a reinforcement interpretation of phenomena such as place learning can be defended in a different but simple way. First of all, reinforcement was responsible for whatever it was that was learned. That this is so can readily be demonstrated. In the simple experiment for example, if we took naive rats, ones not run and reinforced first from the south box, and placed them instead in the north box right from the beginning, they would not

immediately turn left at the choice point at greater than chance frequency. In other words, reinforcement in the early trials from the south was responsible for what was learned except, apparently, that we reinforced, and the rat learned, something that the rat did not do in these early trials! A possible answer seems to be that while originally reinforcing “turning right” we were also unwittingly reinforcing “turning left” as well. The act development and the act genotype-phenotype distinction concepts make it clear how this is possible. The genotype of the act which was originally performed we would argue was something more complex than turning right. When we reinforced what looked like “turning right,” the phenotype performed and observed by us, we in fact reinforced a more complex genotype which included “turning left” as an alternate phenotype which could, because of this reinforcement, henceforth be more readily induced under certain developmental circumstances. The genotype which actually occurred and was reinforced contained a pair of instructions of the sort

if intra and/or extra maze stimuli are on the right, turn right

if intra and/or extra maze stimuli are on the left, turn left.

In training since the first condition held, we observed only that instruction being executed, only that act phenotype, and it took the testing situation to reveal the existence of the alternative instruction, the alternative phenotype. Similarly in the rotation experiment if place learning with respect to the external environment were displayed, we could similarly show that reinforcement was responsible for this implying that the reinforced programme stated

if extra maze stimuli are on the right, turn right

if extra maze stimuli are on the left, turn left.

So the issue is only in part a matter of which cues have acquired control. It is also how acts are

able to behave within limits in an equifinal fashion. The evolutionary analogy suggests that this is possible because it is programmes not simply behaviors which occur and are reinforced. And programmes can include alternative sets of instructions for achieving the same result under different circumstances. Note however that this does not imply that inductive stimuli are interacting with act development in anything other than a naturalistic causal way. Specifically, it does not imply that the organism will henceforth get to the food irrespective of inducing circumstances. As Restle (1965) has shown in such experiments we can readily alter the circumstances enough so that the end is not achieved, so that an incompetent act occurs.

Phenomena analogous to place learning are known to occur in evolution. An organism which has evolved, like an act which has been learned, develops within limits in an equifinal fashion as well. In experimental embryology this is referred to as “regulative development” and was discovered by Hans Driesch (1914) around the turn of the century. In one early experiment for example, Driesch removed one of two cells of a sea urchin embryo at the two-celled stage, and eventually, a whole but small embryo developed anyway! Later research has revealed many such phenomena. For example, with the early limb buds of a salamander:

If half a limb bud is destroyed, the remaining half gives rise to a completely normal limb.

If a limb bud is slit vertically into two or more segments, while remaining an integral part of the embryo, and the parts are prevented from fusing again by inserting a bit of membrane between them, each may develop into a complete limb.

If two limb buds are combined in harmonious orientation with regard to their axes, a single limb develops that is large at first but soon is regulated to normal size.

Berrill, 1976, p. 309)

Here, just as does an act often in the maze cases, under altered circumstances, the direction of

development of an organism structurally re-orient itself, and normal form and/or size results anyway. An example analogous to the behavioral swimming case occurs when a completely different form, but one equally adapted to the circumstances inducing it, results. Some species of plants for example, develop wholly different phenotypes depending upon whether the seed falls on land or in the water. Of course such regulative abilities are not universal. They vary with the taxon, the individual, the stage of develop, and the type of interference. Just as with behaviors, conditions can readily be altered sufficiently so that regulative development fails.

The mechanism by which such limited equifinality or purposefulness in development is executed is not understood to this day. It for long supplied motivation to those who embraced vitalism as a philosophy of the life sciences. Driesch himself was a foremost advocate of such a philosophy, claiming that only the presence of an “entelechy” or life force could account for such apparently purposeful behavior. Few biologists today would take such a position. Although the mechanism of execution is not understood, most would take the performance view that such events are possible ultimately because instruction directing them are encoded in an organism’s genome. In other words, the instructions in a genome must include complex sets of alternative instructions of the sort if X1 is encountered then do Y1, if X2 is encountered then do Y2, and so on. Genomic information must include complex sets and subsets of such instructions. We have here argued essentially the same thing about the limited equifinality of acts displayed in place learning behavior for example. To the extent that an act displays the ability to come up with an altered but equally competent topography when development is interfered with, then such a possibility must have been included as one of the potential developmental pathways in the act “genome” reinforced in the first place. This view is recognizably similar to Hull’s concept of a

“habit family” which he believed to be involved when ‘more than one distinct action sequence (leads) to the attainment of a particular goal or sub-goal’ (1934, p. 40). According to Hull, when one member of such a family is reinforced, “the learning is transferred without specific practice to the remaining members” (1934, p. 41). The phrase “remaining members” though must be carefully interpreted to mean the members of some actually existing set, not any and all actions which would have the same result.

The issue involved in place learning again boils down to whether learning is analogous to a Lamarckian or to a Darwinian evolutionary process. Let us delve more deeply than previously into the problems the Lamarckian view encounters. The main features of Lamarck’s explanation of evolution were that new variants are oriented with respect to adaptation rather than blindly oriented because during development, organisms acquire characteristics that adapt them to their circumstances, and these acquired characteristics are inherited by their offspring. This view of evolutionary theory survived through the introduction of the Darwinian alternative. Eventually it was August Weismann more than any other single person who was responsible for convincing biologists that it was untenable. First of all, Lamarck’s theory contains no satisfactory explanation for adaptations. If we think, good natural scientists that we all are, of the organism-environment interaction purely in causal terms, then why should an environmental event just happen to spark those alterations in an organism which will adapt it to the influence in question? In fact empirically sometimes they do, and sometimes they do not. Organisms after all do often fail to adapt. The same point is true about behaviour. If we think of inductive stimuli as interacting with acts in a naturalistic causal way, why should inducers affect developing acts in just the way required for reinforcement to be achieved? The fact again is that sometimes they do

and sometimes they do not. Secondly, in evolution, adaptations which are purely somatically acquired are not inherited, i.e. the same environmental influence which would cause a phenotype to take on a particular form (say the callouses on a human hand induced by manual labour) would not cause an alteration in the genome of the subject's sex cells such that their offspring would possess the same modification, hand callouses. Weismann once amputated the tails of a number of mice for several generations to drive home the point that their offspring did not thereby end up with shorter tails! I believe that it is also within the realm of possibility to experimentally establish this second point of the "anti-Lamarckian" argument about acts in learning. The same inductive stimulus which alters the phenotype of an act in a particular way would not be likely to alter the neurons involved or the firing patterns of these comprising the act's "genotype" itself such that on its next occurrence the act would be likely to have a similarly altered phenotype. If a topography or phenotype is inductively "forced" on an act, this alteration will not be present on the next occurrence of the act unless it is again then inductively forced on it. Speaking generally and not considering in detail the technical problems involved, this should be experimentally demonstratable.

Weismann's solution to the problem of adaptations acquired in the life history of an organism was the preformationist one accepted since that time. An organism will not acquire during its life history adaptations that its genotype did not already specifically equip or at least permit it to acquire in the first place. Such an alternate phenotype adapted to the influence which induces it presumably is present in some cases because that inductive environment has been present at times in the organisms ancestral past and has therefore been one of the sets of circumstances that the genotype has evolved to be able to cope with. In other cases its presence

may be a purely fortuitous pre-adaptation. I would argue the same thing about the competence of acts. To the extent that acts are regulative in their execution that ability exists because it was already a part of the act's genotype to begin with. Most of the acts which occur probably represent act genotypes with incredibly complex sets and subsets of alternative phenotypes possible from each, depending upon how the development of the act is steered by inducing stimuli. But this is far from saying that inductive stimuli are not interacting with behavior in a wholly naturalistic causal way. If we alter the circumstances sufficiently we certainly do get incompetent phenotypes. And it is certainly far from saying that reinforcement is not responsible for learning.

We have organized this discussion around experiments of the kind introduced by Tolman and then round how the theory proposed here would explain them rather than proceeding immediately to how Tolman or other cognitive theorists would explain them. We did this because these theorists have never clearly spelled out the issue involved. Some features are common to most of their formulations. They tend for example to talk about "what is learned" in cognitive terms as "expectations," Tolman's own term was "a sign-gestalt expectation," and there are others. A linguist might say "a rule" or a "deep structure," an anthropologist "a meaning," a sociologist a "norm," and so on. They also tend (as do most other learning theorists) to talk about stimulus control operations or induction in cognitive terms as "cues" or "warnings." In the context of the theory of the learning process suggested here, this cognitive language is perfectly appropriate. Act genotypes really do constitute "expectations." Since a phenotype is produced by a particular combination of genotype and inductive environment, past reinforcement history has provided a genotype which will develop in a competent direction if the same

(expected) inducing environment is encountered. Indeed, biologists too have readily come to use a cognitive language about organismic genotypes, they represent “information” or “programmes” they say. In fact, to the extent that either type of “genotype” is constituted so as to eventually result in positive selection or reinforcement in the face of a number of alternative inducing environments, we would even be justified in referring to either as a “purpose.” I should also add that this cognitive language of expectations and purpose does not make much sense in the framework of a connectionist reinforcement theory of the learning process, one which ignores act development, the genotype-phenotype distinction, and the possibility then of acts developing regulatively to a degree, and which instead imagines “bonds” being created or strengthened between a stimulus and a response. The debate over the cognitive problem has often been put as one between a stimulus-response theory of what is learned, and a stimulus-reinforcer theory of what is learned. In the theory presented here, the “relationship” between inducing and reinforcing stimuli, the “expectation” which is learned is something very real, an act “genotype,” and it has been given a substantive neurophysiological interpretation. In all of these ways then the theory proposed here is definitely a cognitive one but is rare in being at the same time a reinforcement theory.

Where however I must take strong issue with cognitive learning theory is in its tendency to imply by this cognitive language the existence of purposiveness in a more common sense way. Essentially, it has to be said because this is what it boils down to, its tendency to imply a free will explanation for what is going on. The issue is simply an analogue of the Lamarckian-Darwinian one. Do act genotypes represent purposeful wants, desires, or strivings (conscious or not) for reinforcement such that subsequently the normally causal texture of the environment is

suspended and stimulus control events become then “signals” about what this developing act as a microcomponent of a free will should do in shifting its structure in order to achieve the desired reinforcement? A theory of learning of this type would be a full Lamarckian analogue, i.e. it would maintain that during their development acts miraculously acquire just the form required for competence, that this alteration is “inherited” by subsequent acts and so on and this is how through time, the distribution of acts shifts in the direction of competence. Or instead, is the normally causal texture of the environment maintained in the interaction between developing acts and inductive stimuli such that the latter are interacting with the former in a naturalistic causal way steering the course of act development? According to this view, if acts display a degree of nonrandom ability to adjust competently to inductive alterations, it would be because this ability was already a part of the act’s genotype to begin with, either because of prior reinforcement history or fortuitously. The issue in the two realms is the same. Is adaptation achieved because it is what is sought or is it achieved because natural selection controls what evolves? Is competence achieved because it is what is sought or is it achieved because reinforcement controls what is learned? With respect to the cognitive problem, the issue is no different because a genotype of either type is shown to be able to develop more than one phenotype, each adapted or competent with respect to the circumstances which induce it.

Now describing most versions of cognitive learning theory as explaining learning by using microcomponents of a free will implicitly may seem unfair, yet I do believe that this is what they most often boil down to on close inspection. Comparing them with reinforcement theories including the one proposed here is difficult because the two often diverge on a number of points. First, they typically embrace (ad id Tolman) a classical conditioning or associationist

theory of learning, when two or more stimuli bear a particular temporal relationship to each other, that temporal relationship is learned. Note that frequently in these theories it is not implied, in fact it is often specifically denied that such association lead to transfers of stimulus control over behavior as in the Pavlovian conception where through association, the bell comes to elicit salivation. So transfer of control over behavior does not necessarily occur, rather the organism is simply conceived of as having learned and as now “knowing” what will be followed by what. Guthrie (1935, p. 172) once accused Tolman of leaving his rat “buried in thought.” It knows all of the stimulus relationships to which it has been exposed but what determines what it does? At this point such theories often draw the old knowledge/will distinction of scholastic philosophy. Since rewards are demonstrably effective in altering behavior, this is where they are held to come in. What the animal does in response to a cue depends upon whether it wants to experience or avoid the second “what” in the “what leads to what” that it knows. The following examples have been selected almost at random from the multitude which could be cited.

Bitterman (citing Tolman) has described cognitive learning theory as follows:

reward does not affect learning but it itself ‘learned about’ and therefore comes to be ‘anticipated,’ as food which follows the sounding of a metronome may be said to be anticipated by a Pavlovian dog: ... with level of performance determined by the hedonic value of the anticipated consequences. (1975, p. 700-701)

Menzel has characterized Tolman’s view this way:

Tolman held that, in the course of learning, living organisms do not acquire stereotyped responses but, instead, “cognitive maps,” “that this leads to that,” and that performance is thereafter determined by these cognitions, and what Bindra calls “incentive-motivation,” that is, what it is that an organism wants at a particular point in time and space. (1978, p. 69)

Bindra has declared that:

The most general principle determining response production seems to be this: An animal will move toward an eliciting stimulus that is of greater appetitive motivational valence than other eliciting stimuli in that situation at that time. (1978, p. 48)

Bowrer has summarized the cognitive approach this way:

In cognitive psychology, the intentional actions of a person are characterized in terms of his carrying out a plan which has a particular goal. That is, once a given motive or goal assumes top priority, the person selects a behavioral plan from memory, that, given the present circumstances, should bring about that goal. (1975, p. 29)

Jenkins has stated that an account of response production in cognitive language might read like this:

The occurrences of certain response-outcome sequences (perhaps those involving surprising outcomes) are represented in memory and stored in relation to the stimulus conditions in which they occurred. When the situation recurs the representations may be retrieved. If on the occasion of the retrieval the action has a positive valence, a control centre calls for the execution of that motor programme that corresponds to the stored representation of the response. (1978, p. 64)

I do not know how else to describe such views except as free will theories, if such a notion can be called a theory. Through association, the animal knows what leads to what in terms of stimuli, and it will do whatever needs to be done in order to acquire or avoid the second “what,” as it wills, apparently. The common sense view that, apart from what it knows, the organism is held to do what it wills is apparent in the use of verbs like “wants” and “selects,” but is simply disguised by phrases such as “the hedonic value of the anticipated consequences,” “greater appetitive motivational valence,” and “positive valence.”

Despite a fundamental disagreement with his theorizing then and that of his latter-day

descendents, I would argue that Tolman contributed a very important experimental fact to behavior theory, the (limited) equifinality or purposeful quality of developing acts and a cognitive language about what is learned (expectations, even “purposes”) and about stimulus control (cues and warnings) which needs to be employed to describe his findings. I would argue however that this fact can be accommodated within an expanded reinforcement theory of learning - not a stimulus-response connectionist one to be sure, but one which includes act development and the genotype-phenotype distinction, just as the similar ability of developing organisms has been accommodated within evolutionary theory. This kind of naturalistic purposiveness which takes the preformationist perspective that to the extent that such adjustments are made, the ability to do so was part of the initial programme however, is very different from a Lamarckian type vitalistic purposivism which assumes that new variations in programmes occur in the direction required for them to be reinforced. This expanded reinforcement theory stands between the two existing extremes of the cognitive theorists and Skinner on the one hand who claim that the ability to achieve consequences is acquired in learning (implying, unless qualified, omniscient and omnipotent organisms), and S-R connectionist theory on the other hand, which implies that only a single (the performed) topography is acquired. Evidence from place learning experiments supports neither extreme (Restle, 1965). Instead, it supports the view that a limited set of alternative topographies are acquired, inducible under different conditions, but that sufficiently altered conditions, requiring sufficiently altered topographies, can readily throw the execution of an act off course resulting in incompetence. This is not to say that we yet know the mechanism by which this limited purposefulness in either realm is executed or works, “developmentally” a problem which Hull

also tried to solve unsuccessfully. Whatever the truth of this aspect of the problem may be, when we discover what it is, we will have a theory not only of how “purposes” (genotypes with more than one adapted or competent phenotype) are acquired, but how they are executed as well. Without claiming, then that all of the problems of learning have already been solved, perhaps not even all of those raised by Tolman, I would argue firmly that the specific fact of acts which behave within limits in an equifinal fashion is not one of these and that it can be encompassed within an amended version of reinforcement theory, within a cognitive reinforcement theory suggested by the evolutionary analogy.

7. Conclusion

For obvious reasons, we have taken some of the suggestions derivable from a careful application of the evolutionary analogy and used them to point down some of the highways of learning theory rather than to explore all of its biways. Much even most of the work remains to be done, particularly in the realm of teasing out fine grained predictions which would distinguish it from alternative approaches available. What has been presented thus far is a series of alterations in conceptual foundations leading to a theory in other words, not one or a few isolated and independently testable research hypotheses. The two are very different. There are many things which this theory cannot predict. It cannot predict what event will reinforce or punish a particular act; nor what stimulus will induce the development of a particular act phenotype of will induce its development in a competent or incompetent direction; nor what act will pop up as an avoidance response when alternatives re punished; nor what alternative developmental pathways are embedded in a reinforced programme. But then neither does evolutionary theory

predict analogous phenomena such as what environmental influence will select for or against a particular organism, or what precise phenotypic effect an environmental influence will have, or what alternatives will be present in a particular population, or what regulative abilities a developing organism will display. It does however provide a theory of the process that is learning, and one which can encompass the relativity of reinforcement, the relativity of reinforcement of the same act under different stimulus conditions, the three most basic known forms of learning including their “negative” sides, avoidance, and some aspects at least of the apparently cognitive quality of learning. The issues with respect to which the analogy is suggestive in short, historically have been, and remain, the most important basic problems in learning theory.

References

- Berrill, N.J. Development. New York: McGraw-Hill Book Company, 1976.
- Bindra, D. How adaptive behavior is produced: A perceptual-motivational alternative to response-reinforcement. The Behavioral and Brain Sciences, 1978, 1, 41-52.
- Bitterman, M.E. The comparative analysis of learning. Science, 1975, 188 (4189), 699-709.
- Bolles, R.C. The avoidance learning problem. In G.H. Bower (Ed.), The psychology of learning and motivation (Vol. 6). New York: Academic Press, 1972.
- Bolles, R.C. The comparative psychology of learning. In G. Bermant (Ed.), Perspectives on animal behavior. Glenview, Illinois: Scott, Foresman & Co., 1973.
- Bower, G.H. Cognitive psychology: An introduction. In W.K. Estes (Ed.), Handbook of learning and cognitive processes. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.
- Brown, P.L. & Jenkins, H.M. Auto-shaping of the pigeon's key-peck. Journal of the Experimental Analysis of Behavior, 1968, 11, 1-8.
- Campbell, D.T. Blind variation and selective retention in creative thought as in other knowledge processes. Psychological Review, 1960, 67, 380-400.
- Campbell, D.T. Evolutionary epistemology. In P.A. Schilpp (Ed.), The philosophy of Karl Popper. Book I. La Salle, Illinois: Open Court, 1974.
- Catania, A.C. The concept of the operant in the analysis of behavior. Behaviorism, 1973, 2, 103-116.
- Catania, A.C. Commentary: Bindra: A perceptual-motivational view of adaptive behavior. The Behavioral and Brain Sciences, 1978, 1, 55-56.
- Catania, A.C. Learning. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1979.

- Crow, J.F., & Kimura, M. An introduction to population genetics theory. New York: Harper & Row, 1970.
- Darwin, C. The origin of species. New York: New American Library, 1958.
- de Villiers, P., & Herrnstein, R.J. Toward a law of response strength. Psychological Bulletin, 1976, 83, 1131-1153.
- de Villiers, P. Choice in concurrent schedules and a quantitative formulation of the law of effect. In W.K. Honig & J.E.R. Staddon (Eds.), Handbook of operant behavior. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1977.
- Dews, P.B. The effect of multiple S periods on responding on a fixed-interval schedule. Journal of the Experimental Analysis of Behavior, 1962, 5, 369-374.
- Dobrzecka, C., Szwejkowska, G., & Konorski, J. Qualitative versus directional cues in two forms of differentiation. Science, 1966, 153 (3731), 81-89.
- Dobzhansky, T. Mankind evolving. Toronto: Bantam Books, 1962.
- Dobzhansky, T. Genetics of the evolutionary process. New York: Columbia University Press, 1970.
- Driesch, H. The history and theory of vitalism. London: Macmillan and Company Limited, 1914.
- Dunham, P. The nature of the reinforcing stimulus. In W.K. Honig & J.E.R. Staddon (Eds.), Handbook of operant behavior. Englewood Cliffs, N.J.: Prentice-Hall, 1977.
- Epstein, R. Commentary: Bindra: A perceptual-motivational view of adaptive behavior. The Behavioral and Brain Sciences, 1978, 1, 41-91.
- Ewens, W.J. Population genetics. London: Methuen, 1969.

- Fantino, E. Aversive control. In J.A. Nevin & G.S. Reynolds (Eds.), The study of behavior.
Glenview: Scott, Foresman & Company, 1973.
- Garcia, J., & Koelling, R.A. Relation of cue to consequence in avoidance learning.
Psychonomic Science, 1966, 4, 123-124.
- Gilbert, R.M. Psychology and biology. The Canadian Psychologist, 1970, 11 (3), 221-238.
- Gilbert, R.M. Variation and selection of behavior. In R.M. Gilbert & J.R. Millenson (Eds.),
Reinforcement: Behavioral analyses. New York Academic Press, 1972.
- Grene, M. Two evolutionary theories I. British Journal for the Philosophy of Science, 1958, 9,
110-127.
- Grillner, S. Locomotion in vertebrates: Central mechanisms and reflex interactions.
Physiological Reviews, 1975, 55 (2), 247-304.
- Guthrie, E.R. The psychology of learning. New York: Harper, 1935.
- Hempel, C.G. Philosophy of natural science. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1966.
- Herrnstein, R.J. Relative and absolute strength of response as a function of frequency of
reinforcement. Journal of the Experimental Analysis of Behavior, 1961, 4, 267-272..
- Herrnstein, R.J. On the law of effect. Journal of the Experimental Analysis of Behavior, 1970,
13, 243-266.
- Herrnstein, R.J. Quantitative hedonism. Journal of Psychiatric Research, 1971, 8, 399-412.
- Herrnstein, R.J. Formal properties of the matching law. Journal of the Experimental Analysis of
Behavior, 1974, 21, 159-164.
- Hilgard, E.R., & Bower, G.H. Theories of learning. Fourth edition, Englewood Cliffs, N.J.:
Prentice-Hall Inc., 1975.

- Hinde, R.A., & Stevenson-Hinde, J. (Eds.). Constraints on learning. New York: Academic Press, 1973.
- Houston, John P. Fundamentals of learning. New York; Academic Press, Inc., 1966.
- Hoyle, G. Origin of behavioral rhythms in insect nervous systems. In J. Salanki (Ed.), Neurobiology of invertebrates: Mechanism of rhythm regulation. Budapest: Akademiai Kiado, 1973.
- Huber, F. Principles of motor co-ordination in cyclically recurring behavior in insects. In P.N.R. Usherwood & D.R. Newth (Eds.), Simple nervous systems: An international symposium. London; Edward Arnold (Publishers) Ltd., 1975.
- Hull, C.L. The concept of the habit-family hierarchy and maze learning: Part I. The Psychological Review, 1934, 41, 33-54.
- Hull, C.L. Principles of behavior. New York: Appleton, 1943.
- Hull, C.L. A behavior system. New Haven: Yale University Press, 1952.
- Jenkins, H.M. Commentary: Bindra: A perceptual-motivational view of adaptive behavior. The Behavioral and Brain Sciences, 1978, 1, 64.
- Kaplan, A. The conduct of inquiry. San Francisco: Chandler Publishing Company, 1964.
- Kimmel, H.D. Instrumental conditioning of autonomically mediated response in human beings. American Psychologist, 1974, 29 (5), 325-335.
- Kitchener, R.F. Behavior and behaviorism. Behaviorism, 1977, 5 (2), 11-71.
- Kohler, W. The mentality of apes. New York: Harcourt Brace, 1925.
- Li, C. Population genetics. Chicago: University of Chicago Press, 1955.
- Mackintosh, N.J. The psychology of animal learning. New York; Academic Press, 1974.

- Mayr, E. Evolution and the diversity of life. Cambridge, Mass.: Belknap Press, 1976.
- Meehl, P.E. On the circularity of the law of effect. Psychological Bulletin, 1950, 47, 52-75.
- Menzel, E.W. Jr. Commentary: Bindra: A perceptual-motivational view of adaptive behavior. The Behavioral and Brain Sciences, 1978, 1, 68-69.
- Mettler, L.E. & Gregg, T.G. Population genetics and evolution. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1969.
- Miller, N.E., & DiCarla, L. Instrumental learning of heart rate changes in curarized rats: Shaping and specificity to discriminative stimulus. Journal of Comparative & Physiological Psychology, 1967, 63, 12-19.
- Mowrer, O.H. On the dual nature of learning -- A re-interpretation of "conditioning" and "problem-solving." Harvard Educational Review, 1947, 17, 102-148.
- Mowrer, O.H. Learning theory and behavior. New York: Wiley, 1960.
- Parnas, I., Armstrong, D., & Strumwasser, F. Prolonged excitatory and inhibitory synaptic modulation of a bursting pacemaker neuron. Journal of Neurophysiology, 1974, 37, 594-608.
- Parnas, I. & Strumwasser, F. Mechanism of long-lasting inhibition of a bursting pacemaker neuron. Journal of Neurophysiology, 1974, 37, 609-620.
- Premack, D. Towards empirical behavior laws. Psychological Review, 1959, 66 219-233.
- Premack, D. Rate differential reinforcement in monkey manipulation. Journal of the Experimental Analysis of Behavior, 1963, 6, 81-89.
- Premack, D. Reinforcement theory. In D. Levine (Ed.), Nebraska symposium on motivation (Vol. 13). Lincoln: University of Nebraska Press, 1965.

- Pringle, J.W.S. On the parallel between learning and evolution, Behavior, 1951, 3, 174-215.
- Rachlin, H. On the tautology of the matching law. Journal of the Experimental Analysis of Behavior, 1971, 15, 249-251.
- Rescorla, R.A., & Holland, P.C. Some behavioral approaches to the study of learning. In M.R. Rosenweig & E.L. Bennett (Eds.), Neural mechanisms of learning and memory. The M.I.T. Press, 1976.
- Rescorla, R.A., & Solomon, R.L. Two-Process learning theory: Relationships between Pavlovian conditioning and instrumental learning. Psychological Review, 1967, 74, 151-182.
- Restle, F. Discrimination of cues in mazes: A resolution of the "place-vs.-response" question. In H. Goldstein, D.L. Krantz & J.D. Rains (Eds.), Controversial issues in learning. New York. Appleton-Century-Crofts, 1965.
- Russell, W.M.S. Evolutionary concepts in behavioral science: IV. General System, 1962, 7, 157-193.
- Schoenfeld, W.N. Editorial: Some old work for modern conditioning theory. Conditional Reflex, 1966, 1, 219-223.
- Schoenfeld, W.N. Problems of modern conditioning theory. Conditional Reflex, 1972, 7, (1), 33-65.
- Schoenfeld, W.N., Cole, B.F., Lang, J., & Mankoff, R. Contingency in behavior theory. In F.J. McGuigan & D.B. Lumsden (Eds.), Contemporary approaches to conditioning and learning. Washington: V.M. Winston, 1973.
- Schwartz, B., Gamzu, E. Pavlovian control of operant behavior. In W.K. Honig & J.E.R.

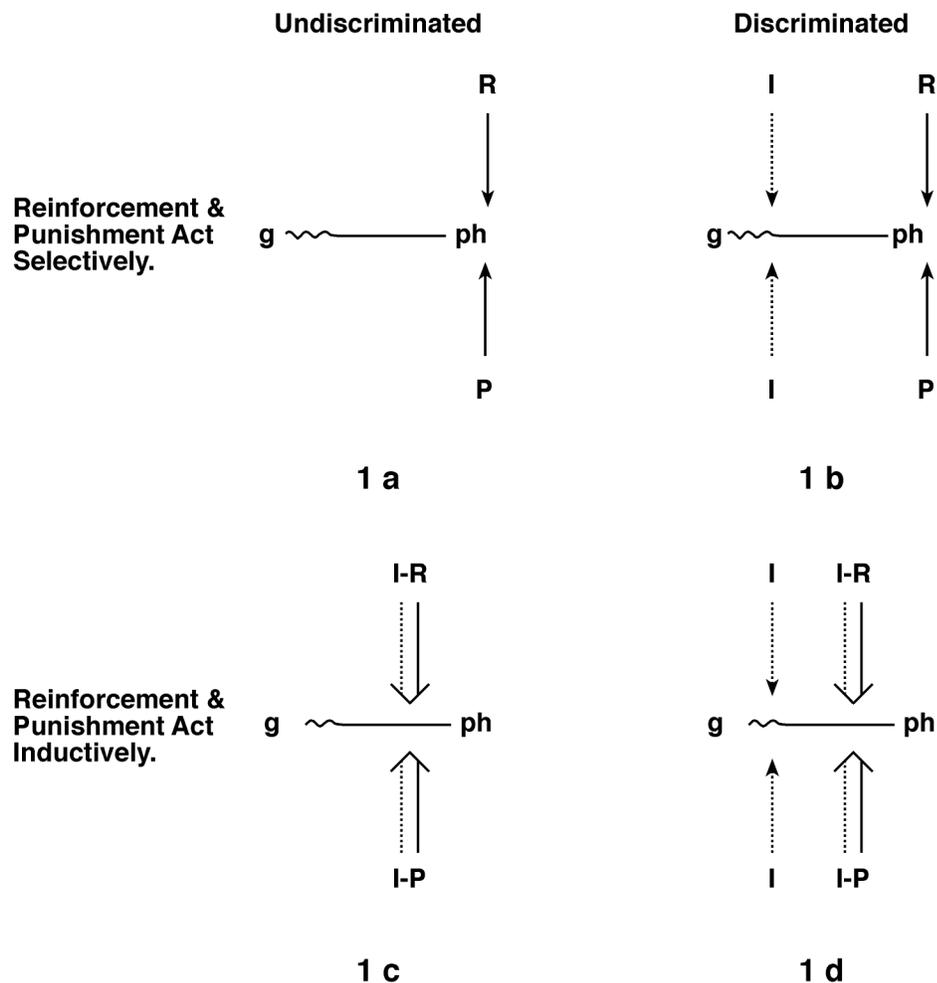
- Staddon (Eds.), Handbook of operant behavior. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1977.
- Seligman, M.E.P., & Hager, J.L. (Eds.)./ Biological boundaries of learning. New York: Appleton-Century-Crofts, 1972.
- Sevenster, P. Motivation and learning in sticklebacks. In D. Ingle (Ed.), The central nervous system and fish behavior. Chicago: University of Chicago Press, 1968.
- Seward, John P. Conditioning theory. In M.H. Marx (Ed.), Learning: Theories. New York: The Macmillan Company, 1970.
- Shettleworth, S.J. Conditioning of domestic chicks to visual and auditory stimuli: control of drinking by visual stimuli and control of conditioned fear by sound. In M.E.P. Seligman & J.L. Hager (Eds.), Biological Boundaries of Learning, New York: Appleton-Century-Crofts, 1972(a), 228-236.
- Shettleworth, S.J. Constraints on learning. In D.S. Lehrman (Ed.), Advances in the Study of Behavior, 1972(b), 4, 1-68.
- Shettleworth, S.J. Food reinforcement and the organization of behavior in Golden Hamsters. In Robert A. Hinde & J. Stevenson-Hinde (Eds.), Constraints on learning. New York: Academic Press, 1973.
- Simon, H. The architecture of complexity. Proceedings of the American Philosophical Society, 1962, 106, (6), 467-482.
- Simpson, George Gaylord. The Meaning of Evolution. New Haven: Yale University Press, 1967.
- Skinner, B.F. The generic nature of the concepts of stimulus and response. Journal of General

- Psychology, 1935, 12, 40-65.
- Skinner, B.F. The phylogeny and ontogeny of behavior. Science, 1966, 153 (3741), 1205-1213.
- Smith, K. Conditioning as an artifact. Psychological Review, 1954, 61 (4), 217-225.
- Staddon, J.E.R., & Simmelhag, V.L. The 'superstition' experiment. Psychological Review, 1971, 78, 3-43.
- Staddon, J.E.R. Learning as adaptation. In W.K. Estes (Ed.), Handbook of learning and cognitive processes. Volume II. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, 1975.
- Strumwasser, F. Neuronal principles organizing periodic behavior. In F.O. Schmitt & F.G. Worden (Eds.), The neurosciences: Third study programme. Cambridge: The M.I.T. Press, 1974.
- Suppe, F. The structure of scientific theories. Urbana: University of Illinois Press, 1977.
- Terrace, H.S. Classical conditioning. In J.A. Nevin & G.S. Reynolds (Eds.), The study of behavior. Glenview, Ill.: Scott, Foresmann & Co., 1973.
- Thorndike, E.L. Animal intelligence. New York; Macmillan, 1911.
- Tolman, E.C., Ritchie, B.F., & Kalish, D. Studies in spatial learning. II Place learning versus response learning. Journal of Experimental Psychology, 1946, 36, 221-226.
- Waddington, C.H. Principles of embryology. London; George Allen and Unwin Ltd., 1956.
- Williams, D.R., & Williams, H. Auto-maintenance in the pigeon: Sustained pecking despite contingent non-reinforcement. Journal of the Experimental Analysis of Behavior, 1969, 12, 511-520.
- Wilson, D.M. Neural operations in anthropod ganglia. In The neuro-sciences: Second study

programme. New York; Rockefeller University Press, 1970.

Zeiler, M. Schedules of reinforcement. In W.K. Honig & J.E.R. Staddon (Eds.), Handbook of operant behavior. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1977.

Figure 1. Relationship among learning procedures as aspects of an evolution-like process*.



*** Key**

g	Analogue of act genotype
ph	Analogue of act phenotype
R and P	(Solid arrows) Reinforcement and punishment
I	(Broken arrows) Environmental induction
I-R and I-P	(Combined broken and solid arrows) Environmentally inductive (i.e.) 'forward-acting' reinforcement and punishment
1 a	Instrumental/operant reinforcement/punishment
1 b	Discriminated forms thereof (the I is an S^D)
1 c	Sensitization/Habituation
1 d	Classical conditioning/conditioned suppression (the I is a CS)