Matching and Melioration in Concurrent Linear VI Schedules

Christopher Cherpas
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Recommended Citation
https://scholarworks.wmich.edu/dissertations/2381

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
MATCHING AND MELIORATION IN CONCURRENT LINEAR VI SCHEDULES

by

Christopher Cherpas

A dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
December 1984
MATCHING AND MELIORATION IN CONCURRENT LINEAR VI SCHEDULES

Christopher Cherpas, Ph.D.
Western Michigan University, 1984

Pigeons were presented with pairs of linear VI schedules in a choice situation. Once a week, the schedules were changed during a special transition session. The steady-state time and response ratios approached matching of reinforcement ratios, although undermatching and bias were prevalent. During the transition sessions, behavioral allocation changed rapidly, in spite of no change in the overall rate of reinforcement; instead, the rate of change in behavioral allocation was found to be an increasing function of differences in local reinforcement rates. The results were interpreted as confirming the theory of melioration (Herrnstein, 1982; Vaughan, 1982b) and disconfirming the theory of economic maximization (Rachlin, 1982).
INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>The Complexity of Reinforcement</td>
<td>1</td>
</tr>
<tr>
<td>The Measurement of Reinforcement</td>
<td>3</td>
</tr>
<tr>
<td>Quantitative Theories of Reinforcement</td>
<td>8</td>
</tr>
<tr>
<td>The Matching Law</td>
<td>8</td>
</tr>
<tr>
<td>Economic Maximization Theory</td>
<td>10</td>
</tr>
<tr>
<td>Melioration</td>
<td>16</td>
</tr>
<tr>
<td>METHOD</td>
<td>24</td>
</tr>
<tr>
<td>Subjects</td>
<td>24</td>
</tr>
<tr>
<td>Apparatus</td>
<td>24</td>
</tr>
<tr>
<td>Procedure</td>
<td>24</td>
</tr>
<tr>
<td>RESULTS</td>
<td>30</td>
</tr>
<tr>
<td>Matching</td>
<td>31</td>
</tr>
<tr>
<td>Melioration</td>
<td>33</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>39</td>
</tr>
<tr>
<td>Melioration Theory</td>
<td>39</td>
</tr>
<tr>
<td>Changeover Responses</td>
<td>44</td>
</tr>
<tr>
<td>Units of Behavior</td>
<td>47</td>
</tr>
<tr>
<td>Reinforcement and Optimal Problem Solving</td>
<td>49</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>54</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1 Concurrent VI Schedule Pairs ......................... 27
LIST OF FIGURES

1 Logged time ratios as a function of logged reinforcement ratios for pigeons 0, 10, and 11. Dashed lines are perfect matching. Solid lines are least-square fits. TL represents time on the right, with TR being time on the right. RL represents reinforcements on the left, with RR being reinforcements on the right. ..................... 32

2 Rate of change in proportional time allocation as a function of differences in local reinforcement rates for pigeons 0, 10, and 11. Data are from transition sessions. .......... 35

3 Rate of change in proportional time allocation as a function of differences in local reinforcement rates for pigeons 0, 10, and 11. Data are from steady-state sessions. ........ 37
INTRODUCTION

The Complexity of Reinforcement

Possibly the single most important variable known to control behavior is reinforcement. Although the powers of "reward" and "punishment" have been described and used for practical purposes for centuries, it is only within the last fifty years that an experimental analysis has been successfully attempted. The Experimental Analysis of Behavior (Skinner, 1957a; Kantor, 1970), which is both the name of a scientific community and a set of methods by which laws of behavior are discovered, is responsible for most of this critical research.

What has been discovered is that reinforcement is a complex phenomenon. Reinforcement "shapes" topographies. It differentiates the intensive and temporal properties of behavior. It also establishes the control of responses by stimuli. Possibly an even more fundamental relation is the control of behavior by reinforcer deprivation (Skinner, 1953). Conversely, the reinforcement effect itself seems to be controlled by the level of reinforcer deprivation (Michael, 1982). As one looks more closely at reinforcement, the list of critical, qualitative issues which define the phenomenon becomes even longer.

1
For example, there are many senses in which one may say that reinforcement is a "relational" principle, but a particularly basic sense has to do with viewing reinforcement as a kind of stimulus-transition, and not simply as an end point. When reinforcement "occurs," a change in the organism's environment takes place. Reinforcement should therefore be defined by both a pre-change and post-change state (Baum, 1974a; Michael, 1975, 1979; Smith, 1974). A further complication is that an analogous response-transition analysis may be required to accurately describe the behavior occurring during reinforcement (Donahoe, 1982; Premack, 1959, 1971).

Narrowing down the scope of analysis to manageable proportions has involved distributing the study of different abstract features of reinforcement across different research programs. Sometimes concentrating on one aspect or another has overshadowed the importance of other factors that are necessary for the total picture. For instance, a distinction has been made between the "shaping" and "strengthening" effects of reinforcement (Morse, 1966), and historically more emphasis has been put on the latter. The same dichotomy has been expressed in terms of
the effects of reinforcement on behavioral change as opposed to its effects on behavioral maintenance, again, the latter having had more research devoted to it.

While particular properties of reinforcement are studied individually, there are efforts to "put the pieces back together" again. Integrative theorizing tends to be of two general types. The first type is the practice of analyzing complex phenomena in qualitative terms, but utilizing the full list of characteristics known about reinforcement. This kind of reinforcement theory is typified by Skinner's book Verbal Behavior (1957b). The other kind of reinforcement theory involves casting one or more functional relations in an equation whose terms represent abstract features of the reinforcement principle. The current study is concerned with this latter kind of theory.

The Measurement of Reinforcement

Before proceeding to an experimental comparison of some current theories of reinforcement, it must be pointed out that those theories originating from within The Experimental Analysis of Behavior tend to be derived from the use of very specific procedures and particular measurements. None of these theories began by explaining the effects of any and
all procedures. Instead, each seems to have gradually evolved toward greater degrees of generality. There is a tendency to incorporate the original concrete measurement of reinforcement into a theory and keep it, even though the theory may eventually interpret a much broader domain; this is true for both the independent (e.g., reinforcement rate) and the dependent variables (e.g., response rate) involved in the process. On the other hand, reinforcement theories that originated outside The Experimental Analysis of Behavior appear to be logically derived from a set of global assumptions. In these cases, the measurement of reinforcement is decided on a more ad hoc basis.

Many measures have been used to assess the strength of reinforced behavior. Skinner (1938, 1969) has consistently promoted the use of response rate. However, one of the major problems of response rate may be that it is itself a conditionable feature of behavior (Nevin, 1974). Average rate of response is an aggregate based on a sequence of interresponse times (IRTs), and Shimp (1969a, 1974, 1982a, 1982b) and others have argued that IRTs are differentially reinforced or "shaped" by conventional reinforcement schedules. This makes response rate more complex than merely the variation of a single, experimentally isolated class of behavior. By relying too heavily on average
response rates, our knowledge of the shaping and strengthening effects of reinforcement may remain unnecessarily limited. On the other hand, there is no doubt that response rate changes dynamically with changes in the conditions of reinforcement, as is convincingly demonstrated in Ferster and Skinner (1957). Another advantage of response rate is that the logical consequences of certain rates of response can be mapped, as feedback functions, for most typical reinforcement schedules (Nevin & Baum, 1980).

Response rate is composed of IRTs or "latencies" from one response to the next (Pennypacker & Johnston, 1980). The term "latency" typically has a more restricted domain, however. The time from the onset of a discriminative stimulus to the occurrence of the first response has been found to vary systematically with parameters of operant reinforcement (Stebbins, 1959; Fath et al., 1983). One of the problems with latency may be that conditions just prior to stimulus onset can affect the situation in complicated ways. Furthermore, the discontinuous nature of a "trial" has been criticized as being rather arbitrary for behavior analysis (Skinner, 1969).

A third measure of reinforcement involves a procedure where a choice of two or more alternatives is available (Herrnstein, 1961). Either the relative number of
responsive or relative time spent at a particular alternative is taken as a measure of how reinforcing a situation is compared to another (de Villiers, 1977). Skinner (1969) and others (e.g., Johnston & Pennypacker, 1980) have opposed this kind of measure, but it has found wide acceptance within The Experimental Analysis of Behavior community. For one thing, a choice situation can reveal differences between parametric values of reinforcement that would lead to a kind of ceiling effect using either rate of response or response latency as the measure of strength. A given behavior can only occur so fast or so soon after a stimulus before it becomes physically impossible to occur any faster or sooner, respectively. This kind of limit is lacking in the choice procedure.

Another advantage is that, conceptually, the choice situation can be interpreted as incorporating the notion that reinforcement is a situation transition (Herrnstein, 1970). With each "choice" or changeover response made from one alternative to the other, an organism experiences a situation transition with procedurally defined pre- and post-change states, each with its own correlated stimulus conditions, response requirements, and reinforcement values. Of course, this arrangement is full of complexities that may not be worth the added sensitivity to parameters (Shimp,
In addition, the notion of behavior as choice is difficult to apply to interactions between multiple schedules (i.e., contrast effects) in which the organism has no control over which schedule component is in effect. In any case, much current research, including the present study, is aimed at explaining concurrent schedule ("choice") performance in terms of particular theories of reinforcement.

There are other measures of reinforcement that have been used successfully. Nevin's (1974) resistance to change in multiple schedules and Hodos' (1963) final completed ratio in progressive ratio schedules are both excellent examples. These diverse measures may eventually be interpreted by a theory which explains reinforcement in terms that are sufficiently primitive and general to cover most situations accurately. Some work has begun to discover equivalencies among these measures (e.g., Fath, et al., 1983).

Meanwhile, there is recent theoretical work that attempts to place reinforcement within a broader, evolutionary framework. While Skinner (1953, 1969, 1974, 1984) has extensively speculated on the nature of reinforcement from the perspective of biological evolution, much current theorizing is carried out in quantitative terms
(e.g., Rachlin, 1982), with the possible advantage that very specific predictions can be tested precisely. The mathematics are generally based on simple economic models. The present study is also concerned with this broader, evolutionary/economic perspective and also takes a quantitative approach to the analysis of reinforcement.

Quantitative Theories of Reinforcement

The Matching Law

In 1970, Herrnstein proposed that the proper measure of strength is the relative frequency of responding in a concurrent schedules situation (Herrnstein, 1970). The relative frequency of response equals or "matches" the relative frequency of reinforcement obtained. That is:

$$\frac{P_1}{P_1 + P_2} = \frac{R_1}{R_1 + R_2}$$

where $P_1$ and $P_2$ are response rates at alternatives 1 and 2, and $R_1$ and $R_2$ are reinforcement rates at alternatives 1 and 2. Both response rates and reinforcement rates are "overall" rates in the sense that the time over which the rates are calculated is the entire duration of an experimental session. Since the denominator for all rate calculations is the same, the terms reduce to simple counts of responses and obtained reinforcements. The matching law
(Equation 1) has been generalized to single schedule and multiple schedule responding when certain parameters are included (Herrnstein, 1970).

Baum (1974b) modified the form of the generalized matching law so that two parameters could be clearly distinguished -- "bias" represented by log k, and "undermatching" or sensitivity represented by a:

$$\log (B_1/B_2) = a \log(r_1/r_2) + \log k. \quad (2)$$

$B_1$ and $B_2$ are behaviors at alternatives 1 and 2 (which may be either in terms of times spent or rates of response), and $r_1$ and $r_2$ are reinforcement rates at the two alternatives. The intercept of the matching line, log k, indicates any consistent degree of preference for one alternative which is controlled by something other than the differential reinforcement supplied by the schedules (e.g., a lower force requirement on one operandum). The "undermatching" parameter, a, is the slope of the matching line. If it is less than one, which it often is (Weardon, 1980), the slope indicates a degree of indifference across the reinforcement alternatives, in spite of obtained differences in reinforcement rates. Exponentiating both sides of Equation 2 yields the familiar power form of the generalized matching law (Baum, 1974b):
\[ \frac{B_1}{B_2} = k \left( \frac{r_1}{r_2} \right)^a. \]  

Regardless of the particular form of the matching law, many have argued (e.g., Shimp, 1982b) that the matching law is not fundamental, but is to be explained at a more molecular level. There is also a theory which states that the matching law is not fundamental for a very different reason, that it is in fact not molar enough.

**Economic Maximization Theory**

Rachlin (1978) presented the theory that matching in concurrent schedules is the outcome of a more fundamental process involving maximizing the total amount of reinforcement (or utility) obtained in an experimental session. Economic maximization theory, as it applies to most concurrent schedules, states that the organism will distribute its choices across the reinforcement alternatives so as to make the sum of obtained reinforcements the highest possible number. Maximization can be stated in simple terms:

\[ RT = C_1 + C_2, \]  

where \( RT \) is the total reinforcements obtained in a session and \( C_1 \) and \( C_2 \) represent the numbers of reinforcements obtained at each alternative.
In concurrent ratio schedules, where a number of responses is required for reinforcement, organisms tend to respond exclusively at the smaller ratio alternative. This is consistent with economic maximization, since responding exclusively on the smaller ratio produces the highest possible rate of reinforcement, and therefore the highest number of total reinforcements per session (RT in Equation 4). Exclusive preference for one alternative is consistent with the matching law, but for a different reason. Exclusive preference involving any set of concurrent schedules automatically produces matching according to the matching law. For example, even if an organism in a concurrent ratios situation spent all of its time at the larger ratio alternative, matching would still result because all of its reinforcements would be obtained at that alternative. Economic maximization theory predicts exclusive preference in concurrent ratios because the total amount of reinforcement is maximized according to specific feedback functions relating responses or time spent at each ratio schedule to the overall rate of reinforcement. The matching law predicts matching in concurrent ratios regardless of the differences in the schedule values and, hence, for trivial reasons.
Concurrent variable-interval (VI) schedules are much more commonly used for studying reinforcement than concurrent ratios. Here, again, economic maximization theory and the matching law predict the same results but for different reasons. The matching law treats the matching of relative behavioral allocation (in either time or responses) to obtained relative reinforcement as a fundamental formulation of the reinforcement principle. It is not reduced to some process which is considered more primitive.

Economic maximization theory, on the other hand, attempts to explain why matching occurs in concurrent variable-interval schedules. Examining the feedback functions, which relate rates of reinforcement to rates of responding (or times allocated) to typical concurrent VI schedules, reveals that an organism "should" approximately match in order to maximize the overall rate of reinforcement in a session (Rachlin, 1982). A problem with this analysis is that the traditional method of programming VI schedules apparently obscures the possibility that some other process is at work which is responsible for matching (Vaughan, 1982b). In fact, one of the significant features of the present study is the use of nontraditional VI schedules. However, the details of this procedure are worth understanding only after one considers the full implications.
and extensions of economic maximization and certain other theories of reinforcement.

Besides concurrent ratios and concurrent VI schedules, situations in which one alternative is an interval (VI) and the other is a ratio (VR) schedule have been studied (Herrnstein & Heyman, 1979; Herrnstein & Loveland, 1975). In such choice situations, matching has been found, with a high degree of bias for the VI schedule (Herrnstein and Heyman, 1979). This is incompatible with economic maximization as stated in simple terms in Equation 4. The problem is that much more time "should" be spent on the ratio side, since a higher combined rate of reinforcement would be obtained by doing so. However, a more general form of economic maximization in concurrent schedules is:

$$ U = C_1 + C_2, \quad (5) $$

where $U$ represents overall "utility," rather than simply overall reinforcement rate (Rachlin, 1982). Utility is broad enough to include both reinforcement and "leisure," where leisure is defined as the opportunity to not work (Green, Kagel, & Battalio, 1982). A ratio schedule, while it may produce a high rate of reinforcement, also induces a high rate of responding -- in other words, a lot of work. Economic maximization theory can account for concurrent VI
VR behavior by assuming that the organism adjusts its
distribution of choices so that an optimal "bundle" of
reinforcement and leisure is obtained in the aggregate. It
just so happens that this distribution coincides with the
findings of biased matching in concurrent VI VR schedules.

Economic maximization theory, a view that an organism's
distribution of behaviors is controlled by the aggregate
utility thereby gained, makes some reasonable explanations
of behavior in concurrent schedules, a standard
proving-ground for current theories of reinforcement.
Economic theory explains a given exclusive preference in
concurrent ratios, while the matching law, by itself, would
be satisfied whether the larger or the smaller ratio is
exclusively preferred. An economic formulation also easily
incorporates the notion of "leisure" into the explanation of
choice.

Additional support for the economic theory is that it
is consistent with the growing branch of experimental
ethology concerned with "optimal foraging." Simply stated,
behaving in an optimal manner means distributing behavior so
that energy intake is maximized while energy expenditure is
minimized (Lea, 1982). During a day of foraging for food,
for example, an organism should spend the least possible
time "handling" or fighting with prey and the greatest
possible time consuming "preferred" prey (be it animal or vegetable). From the perspective of biological evolution, perhaps organisms "should" behave in an optimal or economic manner -- at least approximately. After all, an organism that was behaviorally economic would be more likely to survive and have offspring than one that was not.

Skinner (1984) and others have opposed the notion that evolution has endowed those species whose behaviors are sensitive to consequences with a built-in tendency to "optimize" in terms of energy exchanges with the environment. If animals appear to maximize overall goods and minimize overall costs in natural environments, perhaps it is only because of a particular arrangement of contingencies. Reinforcement, itself, may be seen as operating at a relatively local level, depending primarily on temporal contiguity to bind organismic and environmental events. It need not involve a global, inherent tendency to behave in an optimal manner (Skinner, 1984). Reinforcement does contribute to adaptation and survival under conditions commonly found in nature -- and this probably explains how it was selected as a biological process -- but any theory of "overall adaptation" apparently fails to accurately describe the reinforcement process itself.
According to a "local" view, reinforcement rather arbitrarily binds (i.e., functionally relates) specific stimulus, response, and deprivation elements that occur within close temporal proximity of each other. Far from being universally adaptive, reinforcement can easily be arranged to produce "superstitious" behavior (Skinner, 1948) and "stereotypic" behavior (Schwartz, 1980, 1982). The distribution of behavior found in choice situations is therefore due to the reinforcement of fairly local stimulus and/or response relations, and is not a function of a singular, integrated strategy or pattern carried out by the organism.

Melioration

In recent years, a quantitative formulation of behavior in a choice situation has been developed which, to some degree, incorporates the notion of locally organized reinforced classes of behavior. Termed "melioration," the theory contrasts distinctly with a view of behavior as necessarily integrated and adaptive (Herrnstein, 1982). Melioration characterizes a distribution of choices as a set of locally defined classes of behavior -- in a sense, competing subsystems -- which vary in strength according to their respective local reinforcement rates. Further,
melioration states that an organism's distribution of behavior tends to shift toward topographies and stimuli associated with locally higher rates of reinforcement, despite the overall consequences. When the environment changes, behavior "tracks" the changes produced in the local reinforcement of various classes of behavior. When the environment becomes stable, the discrepancy between the local rates of reinforcement for different response classes drops to zero and behavior ceases to change. It is at that point of stability that the matching law describes the current distribution of response classes. Thus, the matching law is seen as a formulation of the maintenance function of reinforcement, whereas melioration corresponds to the repertoire-altering function of reinforcement (Herrnstein, 1982).

The mathematical formulation of melioration is as follows:

$$d \left[ \frac{T_1}{T_1 + T_2} \right] / dt = f \left( \frac{r_1}{T_1} - \frac{r_2}{T_2} \right). \quad (6)$$

Equation 6 represents the rate of change in preference as a function of the difference in local reinforcement rates at two alternatives. $T_1$ and $T_2$ represent the times spent behaving at the two alternatives and $r_1$ and $r_2$ are numbers of reinforcements delivered. Rate of change in preference
is given as a change in relative time spent at alternative 1, over a period of time, t (Vaughan, 1982a). When there is no difference between reinforcing alternatives, there is no behavioral change:

\[ 0 = \frac{r_1}{T_1} - \frac{r_2}{T_2} . \]  

(7)

Rearranging the terms produces a version of Equation 3, the time-based matching law:

\[ \frac{T_1}{T_2} = k \left( \frac{r_1}{r_2} \right)^a , \]  

(8)

where parameters \( k \) and \( a \) equal 1, which they often approach (Baum, 1979; cf. Weardon, 1982).

In concurrent ratio schedules, melioration predicts that behavior should always shift toward the smaller ratio until exclusive preference is reached, since the larger ratio can never offer a higher local rate of reinforcement. As discussed earlier, economic maximization makes the same prediction, but for a different reason: responding exclusively at the smaller ratio delivers the highest possible overall rate of reinforcement. In concurrent VI schedules, melioration predicts matching (Equation 8), because behavior oscillates in a pattern which tracks local reinforcement rates (Equation 6), resulting in both alternatives delivering equal reinforcement rates (Equation 7). Maximization theory claims, at least for traditional VI
schedules, that matching produces the maximum possible reinforcements from the combined sources. In the asymmetrical situation where the choice of schedules is between an interval and a ratio schedule, behavior stabilizes at a point where the local rates of reinforcement are approximately equal, the result predicted by melioration (Herrnstein, 1982). However, the "bias" observed in such asymmetrical situations complicates the analysis, given a melioration or a maximization framework.

To some extent, melioration theory was anticipated by an "Equalization Principle" (Killeen, 1972; Rachlin, 1973), which states that organisms adjust the time they spend responding across concurrent schedules so that equal rates of reinforcement from the alternatives are obtained. Revusky (1963) noted an "interesting relationship" between matching and the equality of response to reinforcement ratios in concurrent schedules. However, neither of these previous statements were elaborated mathematically or generalized to the extent of the theory of melioration presented in Herrnstein and Vaughan (1980), Herrnstein (1982), Vaughan (1982a), and Vaughan (1982b). Furthermore, current melioration research and theory directly address the evolutionary and economic issues surrounding the reinforcement principle. For instance, it has been
speculated that a local reinforcement process is actually more likely to have biologically evolved than an optimization process (Vaughan, 1982b; cf. Skinner, 1984).

While a local process has been emphasized throughout the discussion of melioration, it is evident from Equation 6 that, at least formally, "locally" simply means that reinforcement rates are calculated in terms of particular times spent behaving at given alternatives, rather than in terms of overall session time. This would hardly seem to be much of a breakthrough in the way of fundamentally explaining the matching phenomenon, at least from the point of view of more molecular theorists (e.g., Shimp, 1982; Silberberg & Ziriax, 1982). However, melioration can be derived from momentary principles of conditioning. For example, Vaughan (1982b) has demonstrated that the trial-by-trial respondent conditioning theory of Rescorla and Wagner (1972), if generalized to operant behavior, mathematically implies melioration. Furthermore, melioration is stated qualitatively in terms of momentary changes in response strength, strictly controlled by changing values in discriminative and conditioned reinforcing stimuli (Vaughan, 1982b). In other words, melioration appears to be a step toward quantifying the molecular level view of reinforcement developed by Skinner.
(e.g., Skinner, 1953), although this interpretation does not seem to be shared by Skinner himself (Skinner, 1984).

The present study is designed to meet a number of criteria. First, it is meant to test the accuracy of melioration theory in a situation which is typical for quantitatively studying reinforcement -- in concurrent variable-interval schedules (de Villiers, 1977). By avoiding asymmetrical schedule arrangements (e.g., concurrent VI VR), the contribution of leisure to the distribution of behavior is not required for the interpretation of the results.

Second, this study is intended to show that the matching law can be explained in terms of a more basic process. Since the matching law formalizes only the maintenance or steady-state function of reinforcement, it does not explain why behavior changes. Consider a situation in which concurrent variable-interval schedules, VI 3 m and VI 1.5 m, have been run for some time. Assume that behavior has stabilized and that the ratio of times spent at the two alternatives approximately equals the ratio (1:2) of reinforcements obtained -- an example of matching. If the schedules are reversed (i.e., concurrent VI 1.5 m, VI 3 m), behavior would have to change in order for matching to be recovered (the ratio now being 2:1). This is exactly what
happens in typical matching studies. Unlike previous matching studies, the present experiment makes use of transitional data so that a dynamic process may be studied.

Finally, the current study aims to distinguish between economic maximization and melioration theories. Previous research (e.g., Herrnstein, 1982; Vaughan, 1982b) has involved a variety of procedures which efficiently contrast the opposing interpretations, much in the manner suggested by Platt (1964). In these studies, melioration appears to be the superior theory. A recent review of the literature in which local rates of reinforcement were reported from concurrent schedules experiments also supports melioration (McSweeney, et al., 1983). In addition, it has been shown that even when the leisure factor is controlled, melioration -- not maximization "minus the costs of behavior" -- accurately describes reinforcement (Boelens, 1984). However, direct evidence of dynamic changes in behavior as a function of local reinforcement rates in concurrent VI VI schedules is generally lacking (cf. Hale & Myerson, 1984; Myerson & Hale, 1984).

In short, the purpose of the present experiment is to test the theory of melioration during transitions in concurrent VI schedules, where it should hold true if it is to explain matching during the steady-states of concurrent
VI. Behavior must change with each change in the pair of VI schedules in order for matching to occur. The critical question is whether the rate of such changes is a function of differences in local reinforcement rates.
METHOD

Subjects

Three female White Carneaux pigeons served. All food (mixed grain) was eaten during experimental sessions, which maintained the pigeons at approximately 88% of their free-feeding weights.

Apparatus

Coulbourne Instruments (model E 10-10) modular test cages were used. Each had three keys, their centers being 8 cm apart. Contingencies and data collection were controlled by SuperSKED software (Snapper & Inglis, 1979), running in a PDP-8a computer made by Digital Equipment Corporation.

Procedure

The procedure used here combined features of two typical methods for arranging choices between schedules of reinforcement: the two-key procedure (Herrnstein, 1961) and the changeover-key procedure (Findley, 1958). In the first, two schedules are programmed concurrently for operands which are spatially distinct; for example, the left response key of a pigeon chamber is programmed for one schedule and the right for the other schedule. In the changeover-key
procedure, a response on a "changeover key" changes the discriminative stimulus in the chamber (and its associated schedule) to the "concurrently" available alternative; reinforcement is produced by pecking a separate "schedule" key.

The critical feature of the two-key procedure is that the two schedule alternatives can be made highly discriminable and topographically distinct. In the present experiment, the left key (illuminated yellow) was the operandum for one schedule; the right key (illuminated white) was the operandum for the other schedule. In other words, different colors and different locations distinguished the two schedule alternatives. Given these conditions, it was hoped that pecking on the left side of the chamber would be easily discriminated from pecking on the right side.

The feature of interest from the changeover-key procedure is that the moment of transition from one schedule situation to another can be made salient to both subject and experimenter. In this case, the center key was used as a changeover key; pecking it turned off the discriminative stimulus illuminating the side key on which the pigeon had just been responding and, after a 1 s blackout, turned on the discriminative stimulus at the other side key, to which
the pigeon would direct further responding. Only after the first peck at the switched-to side key would the center key be illuminated again, making a switch back to the other schedule possible. This contingency prevented the pigeon from making a changeover response without ever sampling the schedule to which the switch was made. The 1 s blackout served in place of the changeover delay (COD) usually employed in concurrent schedules experiments (de Villiers, 1977).

The schedules used were "linear" VI schedules (Vaughan, 1982b). Unlike the typical, nonlinear variety, the linear VI does not wait for a reinforcement to be delivered before timing the next interval to reinforcement availability; instead, successive reinforcements are made available on a strictly "by the clock" basis. If two or more intervals have been timed without a reinforcement delivery, those reinforcements are "stored." As a consequence, obtained rates of reinforcement closely approximate programmed rates, even though a subject's response rate may vary widely. What this implies for concurrent linear VI schedules is that a subject may distribute its behavior across the alternatives in almost any pattern without affecting the overall rate of reinforcement obtained.
Throughout the experiment, the overall rate of reinforcement was programmed at 60 per hour. The ratios of reinforcement rates on the two side keys, in the order in which they were presented, were 1:1, 5:1, 1:5, 4:2, 2:4, and 1:1. Table 1 shows the pairs of schedules in the order of their presentation. The intervals composing the linear VI schedules were determined by the Catania-Reynolds formula (1968). The sequence of intervals for either side key was the result of a computer function which chooses items from lists without replacement (Snapper & Inglis, 1979). The concurrent VIs ran independently, but they were both stopped during all blackouts and all reinforcements. Reinforcement consisted of 5 s access to mixed grain.

Table 1
Concurrent VI Schedule Pairs

<table>
<thead>
<tr>
<th>Condition</th>
<th>Left (Yellow) Key</th>
<th>Right (White) Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VI 2 min</td>
<td>VI 2 min</td>
</tr>
<tr>
<td>2</td>
<td>VI 72 s</td>
<td>VI 6 min</td>
</tr>
<tr>
<td>3</td>
<td>VI 6 min</td>
<td>VI 72 s</td>
</tr>
<tr>
<td>4</td>
<td>VI 1.5 min</td>
<td>VI 3 min</td>
</tr>
<tr>
<td>5</td>
<td>VI 3 min</td>
<td>VI 1.5 min</td>
</tr>
<tr>
<td>6</td>
<td>VI 2 min</td>
<td>VI 2 min</td>
</tr>
</tbody>
</table>
Each session was divided into four session quarters. Each session quarter was terminated by the passage of thirty min, followed by the delivery of the next available reinforcement. At the beginning of each session quarter, the two side keys were illuminated -- the left key, yellow, and the right key, white. Once the pigeon pecked either side key, the opposite side key would darken and the center key would come on. At this point, the VIs would begin timing and the regular procedure would ensue (as described above) until the next session quarter or until the end of the session. The color of the center key depended on which side key was illuminated; it was red when the left key was on (i.e., yellow), and green when the right key was on (i.e., white). Thus, except for the beginnings of session quarters, the pigeon had a choice of pecking either the currently illuminated side key for food or pecking the center, changeover, key in order to switch sides.

Sessions were run seven days a week. Normally, the same concurrent schedule pair prevailed throughout all four quarters of a session. However, once a week, a "transition" session was programmed. In a transition session, the first quarter used the same pair of VI schedules that ran in the previous week. The three subsequent quarters used a new pair of VI schedules, although the sum of their
reinforcement rates was always 60 per hour. For example, the schedules for the first week were both linear VI 2 m (a ratio of 1:1). These schedules also operated in the initial quarter of the first transition session. In the subsequent three quarters of that session, the schedules were linear VI 72 s on the left (yellow) side and linear VI 6 m on the right (white) side (a ratio of 5:1). The program made it impossible for "stored" reinforcements from one quarter to be carried into a subsequent quarter.

In summary, several pairs of linear VI schedules were presented to pigeons in a choice situation. Once a week, the schedules were changed in special transition sessions.
RESULTS

Behavior changed rapidly in reaction to changes in the concurrent schedules and stabilized within one or two sessions following a transition session. Systematic changes in the subsequent five days' data were not evident.

The transitions did not disturb the overall rate of reinforcement, which remained close to 60 per hour. The difference between the first and second quarters' overall reinforcement averaged less than 0.5 per hour during transition sessions. In contrast, local reinforcement rates (i.e., reinforcement rates exclusively correlated with time in the presence of either the yellow or white stimulus) changed radically with changes in the schedule pair. The difference between local reinforcement rates increased sharply in the second quarter of each transition session and gradually decreased in the remainder of such sessions as behavior adjusted to the schedule change. The extent of the difference in local reinforcement rates induced by the transitions depended upon how different the new schedule pair was from the previous pair.
Matching

The steady-state data graphed below come from the last session of each condition -- that is, the last session in which a given pair of VI schedules operated throughout the entire session. Figure 1 shows the logged ratios of times spent in the presence of yellow and white as a function of the logged ratios of obtained reinforcements associated with yellow and white. The dashed diagonal line represents perfect matching (slope = 1.0) without bias (intersecting the origin). The open circles are from the first and last conditions, where the ratio of reinforcements was 1:1 in each case.

The slopes of the least-squares time allocation functions (solid lines) are .86, .75, and .96 for pigeons 1, 10, and 11, respectively; the slopes of the functions for ratios of side key pecks (not shown) are .83, .42, and 1.04. Deviations from matching are generally of the undermatching type (Wearden & Burgess, 1982; cf. Baum, 1979), but pigeon 10 shows an unusual degree of undermatching (i.e., slope less than 1.0). Pigeon 10's behavior is exceptional in other respects, as discussed below.

All pigeons display a bias towards the left (yellow) side, indicated in Figure 1 by the upward displacement of the empirical functions from the theoretical matching line.
Figure 1. Logged time ratios as a function of logged reinforcement ratios for pigeons 0, 10, and 11. Dashed lines are perfect matching. Solid lines are least-square fits. $T_L$ represents time on the left, with $T_R$ being time on the right. $R_L$ represents reinforcements on the left, with $R_R$ being reinforcements on the right.
(Baum, 1974b). Extreme bias is shown by pigeon 10. The open circles in Figure 1 reveal a good reproduction of concurrent VI 2 m, VI 2 m data for pigeons 0 and 11. The open circles noticeably diverge for pigeon 10, however, indicating that this subject's bias changed between the first and last condition of the experiment. These two points are spaced even further from each other in the response ratio function for pigeon 10 (not shown).

The source of pigeon 10's "changing bias" is not exactly known, but visual observation revealed a topographic change correlated with the increase in bias. During the second condition, pigeon 10 acquired the tendency to pace in a circle after nonreinforced response bursts on the yellow key. Nonreinforced bursts on the white key, by contrast, were usually followed immediately by a changeover response. This asymmetrical pattern increased such that, by the third week of the experiment, two or more circles per bout in yellow were common, whereas none occurred in the presence of white.

Melioration

The data graphed in Figure 2 are from the transition sessions. Recall that in a transition session, the pair of linear VI schedules changed at the onset of the second
quarter of the session. Figure 2 shows the rate of change in proportional time allocation as a function of the differences in local (i.e., during yellow versus during white) reinforcement rates for pigeons 0, 10, and 11. The rate of change in proportional time allocation is calculated by subtracting the proportion of time (i.e., time in yellow/time in yellow + time in white) in the first quarter of a given transition session from that in the second quarter, the proportion in the second from that in the third, and so on, yielding three such differences per transition session. These differences are then divided by the time over which the changes occurred (Vaughan, 1982a). Thus, the measure represents the rate of change in behavior during individual transition sessions.

The calculation of differences in local reinforcement rates, the abscissa in Figure 2, involves an attempt to estimate the moving averages of local reinforcement rates assumed by melioration theory to dynamically control behavior (Herrnstein, 1982; Vaughan, 1982b). The local reinforcement rates in the first and second quarters of a transition session are averaged, as are those in the second and third, and in the third and fourth; the average rate in white is then subtracted from that in yellow for all three cases. Thus, the rate of change in time allocation from the
Figure 2. Rate of change in proportional time allocation as a function of differences in local reinforcement rates for pigeons 0, 10, and 11. Data are from transition sessions.
first to the second quarters of a transition would be plotted in Figure 2 as a function of the difference in the average of the local reinforcement rates in both of these quarters.

Linear regression functions (solid lines) are fitted to the data in Figure 2 by the method of least-squares. A clearly increasing function is evident for all subjects. The slopes of the functions for pigeons 0, 10, and 11 are in the same rank order as those of the matching functions in Figure 1. In other words, the more sensitive a subject's changes in behavior were to differences in local reinforcement rates during the transition sessions, the closer its distribution of behavior approximated the distribution of reinforcements in steady-state sessions. This finding supports the view that molar matching is the outcome of a dynamic process in which discrepancies in local reinforcement rates determine whether and in what direction behavior changes.

Figure 2 may be compared with Figure 3, which uses the same axes, but plots data from the steady-state, rather than the transition, sessions. Here there is a restricted range of both local reinforcement rate differences and changes in behavior. Apparently, when local reinforcement rates do not change, behavior does not change. It is also apparent from
Figure 3. Rate of change in proportional time allocation as a function of differences in local reinforcement rates for pigeons 0, 10, and 11. Data are from steady-state sessions.
both figures that pigeon 10's data are highly variable and show considerable amounts of time being spent in the presence of the yellow stimulus. The degree of bias toward yellow is reflected in the extent to which the data lie to the left of the zero-difference point on the abscissa of Figure 3 -- the greater the bias, the further the displacement from zero.
DISCUSSION

Melioration Theory

The results of both steady-state and transition sessions support the view that matching in concurrent VI schedules arises from local, not overall, reinforcement rates. If the pigeons were only interested in maximizing overall reinforcement, it is unclear why their behavior changed so systematically in the face of a constant overall reinforcement rate. Furthermore, changes in behavior were directly correlated with differences in local reinforcement rates, a finding consistent with melioration theory. Another point for melioration is that the slopes of the matching functions (Figure 1) were in the same rank order as the slopes of their corresponding melioration functions (Figure 2), implying that matching only occurs to the extent that melioration is operating.

While these results are encouraging for the theory of melioration, problems remain. First, the psychophysical problem of how to calculate a moving average of local rate of reinforcement is yet unsolved. For example, how recent should reinforcements have occurred to be included in a current average? What kind of averaging method most appropriately translates objective local reinforcement rates
into functional local reinforcement rates? Melioration theory does not currently provide answers to these questions. A theory which melioration is destined to eventually confront on this and other issues is the delay reduction hypothesis (Squires & Fantino, 1971; Dunn & Fantino, 1982; Fantino & Davison, 1983). This theory incorporates the view that the distribution of behavior, and hence matching, is controlled by the differences between local rates of reinforcement and the overall average rate of reinforcement. The delay reduction hypothesis combines local and global measures to predict the conditioned reinforcing values of stimulus conditions having a variety of durations and proximal relations to unconditioned reinforcing events.

A minor, but persistent, problem which plagues all theories derived from choice procedures is that no standard procedure exists which ensures matching. The quantitative analysis of matching, from the perspective of melioration or any other theory, continues to be impeded by extraneous variation in the data (Wearden, 1980, 1983). In the present study, pigeon 10 exhibited some unusual behaviors which might have been prevented by putting a minimum IRT contingency on pecking the schedule keys. If the minimum IRT were exceeded, the stimulus illuminating the changeover
key might have been turned off, and the VI timers stopped, until the next schedule key response was emitted. Possibly small ratios on a changeover key (Dunn, 1982) would be superior to the use of a blackout or a COD to functionally separate the choice components. Until these and other minor procedural variables can be standardized, results can only weakly support or refute theories.

As to the shape of the functions in Figure 2, it should be noted that melioration may be construed as being committed to a linear relation between the rate of change in time allocation and the differences in local reinforcement rates. However, the least-squares functions in Figure 2 are fitted simply to show the general trend of the data. In fact, an S-shaped curve has been used as an example in a recent theoretical analysis of melioration by Vaughan (1982a). The S-shaped curve incorporates the notion that strengthening is a negatively accelerated function of difference in reinforcement rates (i.e., there are "diminishing returns"). The absence of a well defined S-shaped curve in Figure 2 may reflect the method of averaging local reinforcement rates, particularly since the calculation used here probably underestimates the more extreme values of the abscissa.
On the other hand, Myerson and Hale (1984) have shown that, at least when a transition is made from concurrent VIs to concurrent VRs, melioration mathematically implies a linear trajectory for behavioral change. This is because the difference in local reinforcement rates between two ratio schedules should always be linear, that is, if local response rates remain constant. The experimental findings of Myerson and Hale (1984) show that behavior changes most rapidly during the period shortly after the change in schedules is introduced. More rapid changes in behavior occurred during this period in the present study; however, this accelerated change was seen to be correlated with a greater difference in local reinforcement rates. Unfortunately, Myerson and Hale do not report obtained local reinforcement rates during the transition from concurrent VIs to concurrent VRs. It is possible that in the early period of the transition, their subjects acquired a higher local rate of response in the smaller ratio component sooner than in the larger ratio component. This would have temporarily created a greater difference in local reinforcement rates than would be obtained some time later in the concurrent VR VR condition.

Besides the psychophysical problem of accurately calculating "functional" local rates of reinforcement
(Timberlake, 1982), the lack of a standard procedure for the study of matching (Vaughan, 1982a), and the debatable shape of the melioration function (Myerson & Hale, 1984), another challenge for melioration theory is the question of whether more molecular processes might underlie melioration. Momentary maximization (Shimp, 1969) is a highly molecular theory of reinforcement. Like economic maximization, momentary maximization starts with the assumption that the organism's behavior tends toward some optimal relation with the prevailing contingencies of reinforcement. Unlike its more molar counterpart, momentary maximization does not assume that the aggregate good is maximized; rather, it states that the response with the momentarily highest probability of reinforcement should always have the highest probability of occurrence. How melioration theory will answer this kind of challenge is a matter for future experimental research and theoretical analysis; but in the meantime, a basic measurement question which the molecular perspective raises is exactly what units of behavior are generated by schedules of reinforcement, particularly in concurrent VIIs where the interest in quantitative functions has somewhat overshadowed the concern for accurately identifying the units which compose such functions.
Skinner (1950) noted that there must be at least three response classes involved in a typical concurrent schedules procedure: the behavior of responding at any operandum, the behavior of switching in one direction, and the behavior of switching in the other direction. Catania (1966) described four classes of behavior: behaving at each of the two schedules, and switching in both directions. To some extent, the procedural arrangement of stimulus and response conditions in the present study may be interpreted as reflecting the four-response class analysis, each class being correlated with a different color stimulus: yellow, white, green, or red. Shimp (1982b) and Silberberg and Ziriax (1982) promote the position that concurrent VI schedules establish numerous functional classes of interchangeover times (ICTs) and that the reinforcement of these units is responsible for the incidental occurrence of matching. Despite the differences in these analyses, the element common to all appears to be the changeover responses.

The body of literature devoted to changeover responses has steadily grown over the last 10 years. Baum (1974a) found that the rate of changeover responses in a "chained concurrent" choice situation increases as the difference
between schedules decreases. Changeover rate was lowest at the extremes of the preference scale, and highest when the ratio of scheduled reinforcements was 1:1. This inverted-U shape has reappeared recently in the data of Dreyfus et al. (1982), only with a different dependent variable: the proportion of reinforcements delivered immediately after a changeover response. This variable, like changeover rate (Baum, 1974a), was found to be greatest when the difference between the schedules in a choice situation was near zero. The parallel in this set of relations shows that changeover responses are differentially reinforced, even though the procedure is not explicitly programmed to do so. This "contingency" may turn out to be the most important factor determining the distribution of behavior in concurrent schedules. It may also be the case that response rate, as a measure of the reinforcement effect, does not have its usual limitations when it is applied to changeover responses; after all, "time allocation" is simply the inverse of the relative rate of changeovers.

However, just as response rate under single schedules can be viewed as the outcome of the shaping of IRTs, rates of changeover responses can be analyzed into distributions of ICTs (interchangeover times). Silberberg and Ziriax (1982) suggest that the critical response-reinforcement
relation in concurrent VI schedules is the function relating ICTs per opportunity to the reinforcements per opportunity for those ICTs. While it is not clear how this relation would have been studied during transitions in concurrent schedules, this function was plotted for data sampled during steady-state periods of the present experiment (not shown). The relation tended to be positive and linear, indicating that the molecular analysis of matching is orderly at the level of ICT distributions. Shimp (1982b) has conducted experiments in which differential reinforcement for ICTs was explicitly manipulated. In that kind of situation, functional control was clearly demonstrated. However, Shimp argues that in the typical concurrent VI situation, the relation between ICTs and their reinforcement is only correlational, and reflects more of a characteristic of the procedure than any law of behavior.

A variety of other molecular properties of changeover behavior (e.g., sequential patterns of ICTs) have yet to be systematically explored, but perhaps Shimp's argument is relevant to a more general point. The concurrent schedules procedure has produced orderly quantitative relations between certain properties of behavior and reinforcement. Even though the changeover response is a key mediating element in such orderly relations, the concurrent schedules
procedure itself does not allow much direct experimenter control over the fine-grain patterning of changeovers. Likewise, the theory of melioration makes sense of both steady-state and transition data within a somewhat molecular framework, but it cannot be specific about what occurs at the exact moment of switching from one stimulus situation to another. The point here is that some features of the reinforcement principle, such as its capacity to shape behavior into molecular units, may be understandable only by employing a variety of procedures. Understanding all of the important features of reinforcement probably involves using a number of different measures. Finally, if reinforcement can be explained by an integrated theory, it must be one which has rules for determining units of behavior, wherever and however they are measured.

Units of Behavior

Skinner (1938) recommended a strategy for identifying functional units of behavior. By systematically observing the changes in behavior correlated with changes in the environment, one could discover "responses classes," defined by their functional bindings with "stimulus classes." The method also calls for inducing orderly variations in the correlations of stimulus and response classes by
experimentally manipulating some third variable, like reinforcement. These functions define "dynamic laws" of behavior. The dynamic laws, in turn, verify the reliability of the functional units of behavior (Schick, 1971). Skinner (1953) also recognized that units of behavior are somewhat artificial and do not always adequately account for the continuity of behavior, but that units are convenient for identifying the major variables that strengthen behavior:

A more useful way of putting it is to say that the elements are strengthened wherever they occur. This leads us to identify the element rather than the response as the unit of behavior. It is a sort of behavioral atom, which may never appear by itself upon any single occasion but is the essential ingredient or component of all observed instances. (p. 94)

Whether molecular, elemental, or atomic, it appears that the components of behavior which are controlled by reinforcement are not likely to be thoroughly defined by a few passes through a single experimental procedure. In the present experiment, functional control was demonstrated by repeatedly shifting a locally higher rate of reinforcement back and forth across two choice alternatives. While it is not clear precisely what features of local reinforcement or of the two response alternatives themselves are responsible for the orderly changes observed in behavior, the results are easily interpreted in qualitative terms: reinforcement strengthens stimulus and response elements that more or less
immediately precede its presentation. Beyond that, it would seem that only a kind of "recursive" definition can lead to a precise specification of what it is that gets "strengthened." In other words, once functional control is demonstrated in some behavioral units, the constituents of those units must be tested for control by reinforcement in other contexts; if these components are susceptible to reinforcement, they should be further analyzed into constituents, which may or may not be found susceptible to reinforcement. Such a process continues until certain base cases are reached wherein either all possible components are found to be affected by reinforcement or other, perhaps constant, factors are discovered; either way, these components can be thought of as multiply determining the composite units observed in any instance.

Reinforcement and Optimal Problem Solving

A theory like economic maximization has a built-in explanation for the continuity of behavior: all components of behavior should tend to adapt to the environment so that the maximum good is obtained. At the other extreme, virtually no continuity is insured by early reinforcement theories that characterized reinforcement as "stamping in" responses which presumably recur without variation or
generalization. Somewhere in between lies the reality of the situation, a world in which operant behavior sometimes intelligently solves the organism's problems in living, while at other times seeming to be the very cause of it's problems. The fact that such extremes do exist is itself an argument against any theory claiming that behavior, or the evolution of species, follows an optimal problem solving strategy. However, it may turn out that viewing natural phenomena as deviations from optimality may be an important heuristic for discovering the true nature of biological adaptation on any level.

Skinner (1969) has analyzed problem solving in terms of operants, which are essentially units of behavior controlled locally by reinforcement. While it recognized that a field of problem solving could exist independently of the domain of naturally occurring behavior, Skinner's analysis was geared toward how problem solving actually occurs in human behavior. His analysis shows that sequences of operants, controlled by reinforcement, account for much of what is called "problem solving" as it occurs in nature. In addition, Skinner (1953, 1957b, 1969) has described a process in which operants are created or otherwise strengthened by the pairing of (usually verbal) stimuli without any obvious reinforcement of a response. Called
"instruction" or sometimes control by "rules," this additional process is apparently responsible for behavior which has had little opportunity to be exposed directly to a reinforcement contingency, but is still likely to occur when the appropriate situation arises. Once the instructed behavior occurs under the appropriate conditions it is likely to be reinforced by either the social or physical environment; such reinforcement then becomes the dominant source of strength for subsequent occurrences of the behavior.

Rules, while they may not be made to be broken, are notoriously weak. Being told that addicting drugs are physically dangerous is not enough to override most drug addicts' histories of finding the drugs to be reinforcing. Much "rule following" continues because a social environment reinforces the behavior described by the rules or else it is quick to supply the rule again when the behavior appears weak. Because instruction is weak without continuing support from contingencies, it cannot effectively sustain behavior which extends beyond or is in competition with local reinforcement control. Thus, the long term aversive consequences for overeating, smoking, and drinking alcohol, as well as the rules that describe these consequences, are often easily counteracted by the immediate reinforcement for
engaging in these behaviors. Countless examples exist in which it is obvious that the local reinforcement for some behavior prevents the behaver from obtaining some global good.

One conclusion which this perspective forces is that to achieve long term goals, to behave as an "optimal problem solver," requires an environment designed specifically to deal with the local, "nonoptimal" organization of behavior produced by reinforcement. The recent development of sophisticated problem solving machines, computers, could surely aid in structuring local contingencies under which certain long term benefits could be maximized. On the other hand, a "Skinner box" existence may not sound very appealing to the typical person at present.

Another conclusion which is suggested is that the traditional concept of a "self" with a single, global set of motivations or "will" is probably an inaccurate description of the human condition. As Skinner (1953, 1971, 1974) has pointed out, we must have multiple "selves," each corresponding to a set of reinforcement contingencies. A careful reading of these works reveals that many cultural issues are critically concerned with the control of behavior by local reinforcement, although such issues may not be generally recognized as such.
A final conclusion of this discussion is related to scientific methodology. Assuming that scientific behavior is also controlled by local reinforcement, one would not expect that the conduct of science is actually directed by so-called "-isms," such as empiricism, realism, or even behaviorism. To improve scientific behavior, which is the supposed aim of scientific methodologies, requires an understanding of scientific behavior in terms of local reinforcement contingencies. Such a "task analysis" could be used as a guide for teaching scientific behavior and as a basis for developing better tools for facilitating quality performances. A detailed behavioral analysis would also aid in the creation of an "expert" computer system which mimics the behavior of scientific experts. By combining general problem solving techniques (being developed by members of the artificial intelligence community) with a local behavioral analysis of scientific behavior, contingencies could be constructed which would help scientific behavior to approach a kind of "optimal" performance.


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


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


