INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

I. The Dynamics of Operant Conditioning: Theory and Experiments. II. Context-Dependent Form Perception: Theory and Psychophysics

by

Valentin Dragoi

Department of Psychology: Experimental
Duke University

Date 5-16-97

Approved:

John E. R. Staddon, Supervisor

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor in Philosophy in the Department of Psychology: Experimental, in the Graduate School of Duke University, 1997.

ABSTRACT

I. The Dynamics of Operant Conditioning: Theory and Experiments. II. Context-Dependent Form Perception: Theory and Psychophysics

by

Valentin Dragoi
Department of Psychology: Experimental
Duke University

Date 5-16-97

Approved:

John E. R. Staddon, Supervisor

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor in Philosophy in the Department of Psychology: Experimental, in the Graduate School of Duke University, 1997.

Abstract

My thesis describes results in two areas: operant conditioning and vision.

Operant conditioning. It has been traditionally assumed that operant conditioning is either driven by events in the recent past (local or short-term memory processes) or by events in the more remote past (global or long-term memory processes). These views have been extremely influential in the operant conditioning community and have imposed severe constraints on the experimental and theoretical analysis of behavior for over 50 years. However, increasing evidence suggests that neither local nor global models provide a satisfactory description of all data. Local models lack the capacity to encode events in the remote past, whereas global models are relatively insensitive at short time scales. By combining experimental and modelling techniques I propose here an alternative view which claims that learning is driven by *both* local and global processes. Based on interactions between short- and long-term memory mechanisms I propose a real-time model that explains the major static and dynamic properties of operant behavior in both single-choice and multi-response situations. The model is utilized to advance new predictions by devising a set of theory-guided experiments that investigate how the combined effect of long-term training history and reinforcement probability affects operant behavior.

Vision. Spatial vision is context dependent, i.e., the perceived visual attributes of a target stimulus depend on the context within which the target is placed. Geometrical illusions, which are context-induced subjective distortions of visual features, such as length, orientation, or curvature of lines, are the most striking example. Physiological studies that use stimulus configurations similar to those used in the psychophysics of visual illusions have demonstrated that responses of orientation-selective cells in visual cortex can be suppressed or facilitated in the presence of an

oriented visual context in the same direction as the perception of lines is facilitated or suppressed by surrounding line elements. To understand the fundamental properties of context-dependency at the physiological and perceptual level I use a combined modelling and psychophysical approach with which to investigate surround-dependent effects in visual cortex, as well as psychophysical phenomena such as geometrical illusions of orientation and extent. I propose a model of visual processing in the primary visual cortex that reconciles conflicting experimental data reporting both suppressive and facilitatory orientation-dependent context effects. The model suggests a new role for intracortical connections and cortical feedback projections and suggests explanations for the genesis of geometrical illusions of orientation and extent. Psychophysical experiments are used to investigate the nature of the Müller-Lyer geometrical illusion, one of the best-known and most extensively investigated geometrical illusions, by showing striking correlations between the perceived length distortion in the Müller-Lyer illusion and 'low level' visual processing such as detectability of a luminance bar (target), in stimulus configurations that contain a single set of inward and outward-pointing arrowheads.

ACKNOWLEDGEMENTS

My deepest respect to John Staddon, whose charismatic influence was crucial for making me understand what love and respect for science really are. To Greg Lockhead, who provided much inspiration, encouragement, and advise during the past four years, I owe my huge gratitude. Many thanks to Nestor Schmajuk who helped a great deal especially when his ideas were in opposition with my own views. David Fitzpatrick taught me how to ask cogent questions about the properties of visual cortical circuitry that are responsible for cognitive phenomena. I also thank Dale Purves, Richard Palmer, and Larry Katz for many insightful discussions at different points during my PhD research. The interaction with Mark Cleaveland, Silvano Zanutto, Armando Machado, Jose Bentes, Catalin Buhusi, and the remaining L. A. B. members, was extremely beneficial for organizing my scientific ideas and for preparing me better for life.

My huge gratitude to you, Ioana, Tudor, George, and parents, for showing me the real values in life.

TABLE OF CONTENTS

LIST OF FIGURES

The Dynamics of Operant Conditioning

Abstract

Existing models for free-operant and discrete-trial instrumental learning are relatively insensitive to historical properties of behavior and applicable to only limited data sets. I propose a minimal set of principles based on short- and long-term memory mechanisms that can explain the major static and dynamic properties of operant behavior in both single-choice and multi-response situations. The critical features of the theory are: (a) that the key property of conditioning is assessment of the *degree of association* between responses and reinforcement and stimuli and reinforcement; (b) that contingent reinforcement is represented by *learning expectancy*, which is the combined prediction of response-reinforcement and stimulus-reinforcement associations; (c) that the operant response is controlled by the interplay between facilitatory and suppressive variables which integrate mismatches between expected (long-term) and experienced (short-term) events; (d) that very long-term effects are encoded by a *consolidation memory*. Learning dynamics in this model are sensitive to the entire reinforcement history. The model predicts the qualitative features of operant phenomena such as response selection, contingency effects, effects of reinforcement delay, matching in choice experiments, development of preference, contrast effects, resistance to extinction, spontaneous recovery, regression, serial-reversal learning, and the overtraining reversal effect. The theory offers a small set of elementary principles that may help resolve the long-standing debate about the fundamental variables controlling operant conditioning. Experimental tests that covary reinforcement probability and length of training are devised to understand the genesis of choice behavior.

Introduction

Theories of operant learning have traditionally emphasized static (asymptotic) principles: laws for stable equilibria that are largely independent of the organism's previous history. The best-known example is the *matching law* (Herrnstein, 1961), according to which the proportion of responses to one choice alternative matches the proportion of reinforcements delivered by it. Equilibrium laws are useful, but do not take us very far towards understanding the underlying causal mechanism.

In contrast, dynamic theories specify how operant behavior is shaped by reinforcement in real time (or response-by-response). One of the most influential concepts in dynamic theories for operant conditioning is the "leaky integrator" (Bush & Mosteller, 1955), a linear operator whose output, response strength, increases following each occurrence of an external stimulus (e.g., a reinforcement) and declines with the lapse of time in the absence of stimulation. The idea of leaky integrator is incorporated in virtually every dynamic model of operant conditioning (e.g., Luce, 1959; Myerson & Miezin, 1980; Vaughan, 1982; Lea & Dow, 1984; Staddon & Zhang, 1991; Davis et al., 1993; Killeen, 1994). All these models have in common the hypothesis that the time constant of the integrator (which defines the temporal window of event integration) is fixed. A small time constant means high sensitivity to current events, but the effects of each event persist only for a short time; whereas a large time constant means low sensitivity to current events, but effects persist longer. Most current theories of operant learning are exclusively local (short-term) or global (long-term). With few exceptions (e.g., Davis et al., 1993), local processes have been favored.

Local dynamic theories of operant conditioning predict behavior is on the basis of events (responses and reinforcements) in the recent past. Much can be explained with short-term (history-independent) models. For instance, a local model is sufficient to explain effects which occur within one experimental session, such as PREE (e.g., Kacelnik et al., 1987), effects of reinforcement probability ratio and absolute difference on the development of choice behavior (Mazur & Ratti, 1991; Mazur,

1992).

However, operant conditioning is often *history dependent*: the same set of experimental conditions lead to different performances, depending on earlier conditions. Short-term dynamic models of conditioning cannot explain data showing that current performance is affected by remote reinforcement history. For instance, responding to a previously reinforced stimulus recovers after prolonged extinction (spontaneous recovery: see Robbins, 1991, for a review), the ability to perform reversals improves with increasing training (serial reversal learning: e.g., Davis & Staddon, 1990), and the partial reinforcement extinction effect (more resistance to extinction following partial reinforcement) is reversed (more resistance to extinction following continuous reinforcement) if training is extensive (reverse PREE: Nevin, 1988). Apparently, any realistic dynamic model must take into account events in the relatively remote past.

A problem with many global models is that they are relatively insensitive to events at short time scales. For example, there is recent evidence (Dreyfus, 1991; Mark & Gallistel, 1994) that in concurrent VI-VI schedules animals are able to track local variations in relative rate of reinforcement and to attain matching more rapidly than previously thought. These results do not seem to be explicable by any process that relies exclusively on long-term dynamics[1].

Since neither local nor global dynamic models provide a satisfactory description of all data, we hypothesize that learning is driven by *both* local and global processes. I here propose a real-time model of operant conditioning that involves processes at several time scales to explain the major static (asymptotic) and dynamic (transient) properties of reinforcement learning in animals. Consistent with the views of Tolman (1932), as well as with most of the classical conditioning theorists (e.g., Rescorla & Wagner, 1972; Grossberg, 1982; Daly & Daly, 1982; Sutton & Barto, 1981, 1990; Klopf, 1988; Schmajuk & DiCarlo, 1992; Schmajuk,

[1] In fact, given a suitable nonlinear response rule, such as winner-take-all, even a very long-term theory, such as the cumulative-effects (CE) model of Davis, et al., 1993, can explain results like those of Mark & Gallistel. But this model has other limitations, as we show below.

in press), I hypothesize that the organism generates expectancies or predictions of future events based on the experienced reinforcement. However, unlike existing theories, we analyze the dynamics of conditioning by defining two expectancies: a short and a long-term expectancy (a third, long-term, process is also needed to account for a few effects). I describe the essential properties of acquisition and extinction in real-time operant behavior in terms of the interplay between these two expectancies (previous versions of the theory were presented in Dragoi, in press, and Dragoi & Staddon, 1993).

Overview

The focus here is on operant behavior in *transition*, i.e., the paths through which equilibria (stable states) are reached. I look at *response rate* as dependent variable, leaving temporal discrimination for later development. (This can be justified, because some species that show the major phenomena of operant learning, show poor, or no, temporal discrimination: e.g., goldfish, Rozin, 1965). I consider the length of training (number of reinforcements or extinction trials) as a key variable which distinguishes between two classes of phenomena with different time courses. The dynamics of responding are analyzed in the following situations: (a) *Effects that are defined within a few experimental sessions*: assignment of credit (response selection, delayed reinforcement, delayed reinforcement and preference reversal, the effect of noncontingent reinforcement); development of preference in choice behavior (effects of ratio and absolute difference between the reinforcement probabilities); matching; successive contrast effects; behavioral contrast effects; partial reinforcement extinction effect; overtraining reversal effect; effects of context on stimulus preference. (b) *Effects that occur only after extended training*: effect of length of training on resistance to extinction ("reversed PREE"); spontaneous recovery and regression; serial reversal learning (successive daily reversals and reversals in blocks of days).

The theory is based on the following behavioral principles:

(1) *Response competition.* Operant response units mutually inhibit each other such that the stronger response will have a higher probability of occurring.

(2) *Short and long-term memory traces.* Each response leaves a brief, decaying short-term memory trace. Associations between responses and the reinforcement leave both decaying short and long-term memory traces. The associations increase in strength with the degree of contiguity between responses and the reinforcement.

(3) *Learning expectancy.* The organism builds short- and long-term reinforcement expectancies based on recent and remote association memory. Short-term reinforcement expectancy is defined as the product of response- and short-term memory traces. Long-term reinforcement expectancy is defined as the product of response- and the long-term memory traces. It is proposed a product rule here because some kind of nonlinear combination seems to be necessary to solve the assignment-of-credit problem (cf. Staddon & Zhang, 1991, for a similar proposal).

(4) *Expectancy mismatch drives the operant response.* Long- and short-term expectancies are compared to detect deviations from learned contingency relations. Their difference controls the operant response. If short-term expectancy is greater than long-term expectancy (reinforcement is underpredicted, reinforcement conditions are improving – if reinforcement increases in magnitude or probability, for example) the strength of the operant response increases. This is termed *behavioral excitation.* But if short-term expectancy is greater than long-term expectancy (reinforcement is overpredicted, e.g., reinforcement decreases in magnitude or probability) the strength of the operant response decreases. This is termed *behavioral inhibition.* The amount of behavioral excitation or inhibition is proportional to the expectancy mismatch.

(5) *Consolidation long-term memory.* Short and long-term memory variables are both *transient*: even long-term memory declines to zero eventually. To represent *permanent* changes, the model includes a *consolidation long-term memory*, which increases via slow changes in the association between behavioral excitation and the response. Consolidation long-term memory is effective in situations in

which training is extended, allowing for differences between an "experienced" and a "naive" animal.

Similar principles apply to stimuli and responses[2]. Phenomena defined within a few experimental sessions using the interaction between principles (1)-(4), short and long time scales. Phenomena that occur only after extended training require principle (5): consolidation long-term memory, which has a time scale sufficient to capture the entire reinforcement history. Consolidation long-term memory is not necessary to explain transient effects, but it does not impair predictions derived from the other assumptions and is necessary to explain effects that depend on extended training.

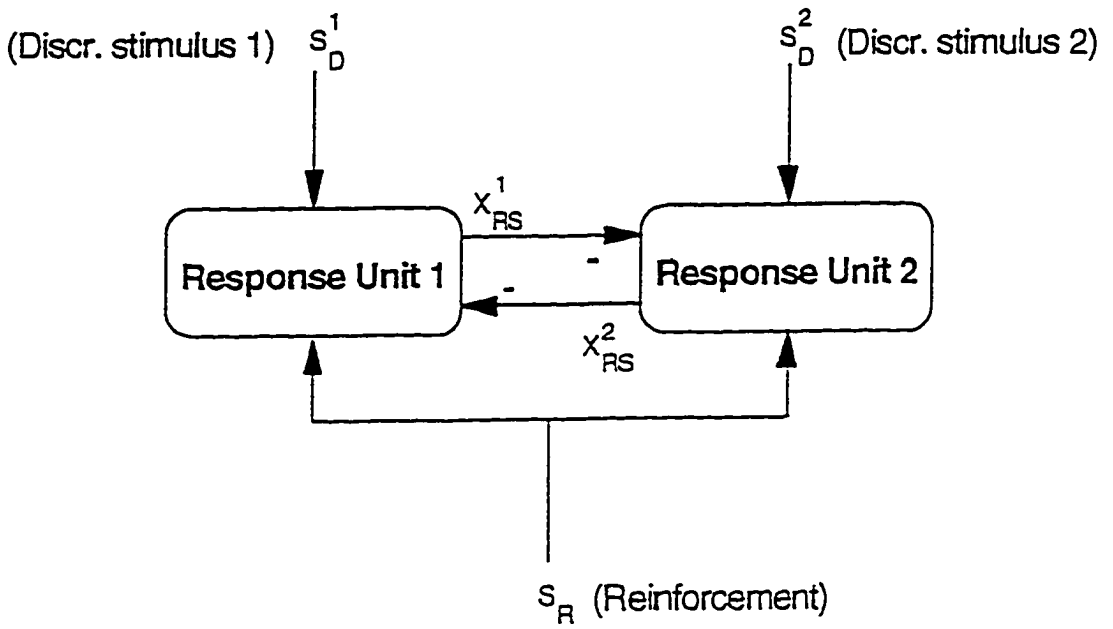
The model

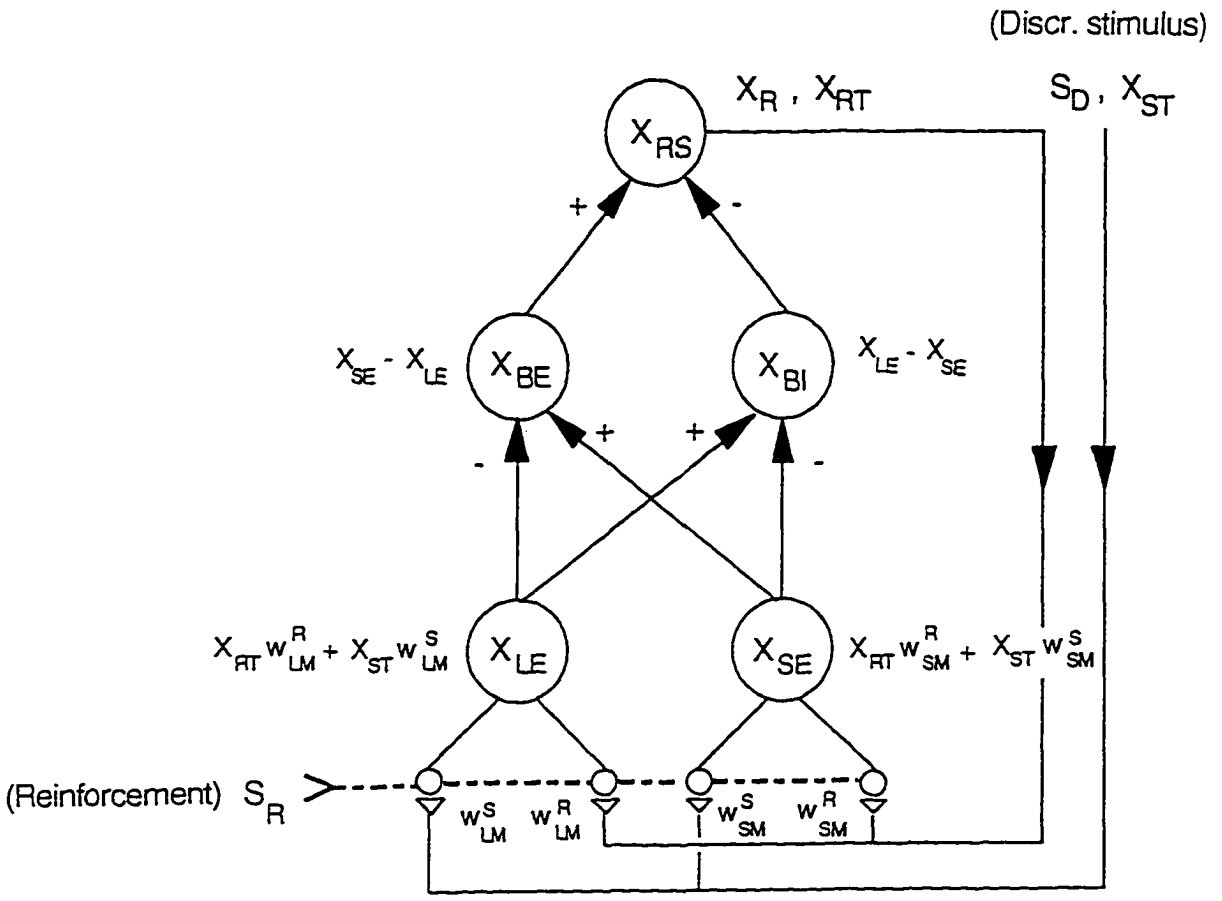
Figures 1 and 2 show the model's structure and dynamic behavior. Figure 1A illustrates the overall structure of the model with two mutually inhibitory response units. Operant response units use their response strength outputs (X^1_{RS} and X^2_{RS}) to inhibit each other to ensure that the response with the higher associative strength will be selected by the reinforcement. Notice that each response unit has its own internal structure and is controlled by a specific discriminative stimulus (S^1_D or S^2_D), the only element shared with other units being the reinforcing stimulus (S_R).

Figure 1B (the response-unit diagram) shows the details of one response unit. Figure 2 illustrates the behavior of model variables. Responses (X_R), defined in the conventional way as lever presses or key pecks, reinforcers (S_R), and discriminative stimuli (S_D), are binary events, i.e., they are equal to 1 when a response is generated, a unit of reinforcement is awarded, or the discriminative stimulus is turned "ON", and they are 0 otherwise.

[2]Stimuli and responses are treated completely symmetrically such that the assignment-of-credit analysis (below) works equally well for both.

Figure 1 (A) *Model diagram for two interconnected response units.* Each response in controlled by a different discriminative stimulus. The reinforcement is common to both response units. Response competition is implemented by mutually inhibitory connections. (B) *Detailed diagram of the model for one response unit.* X_{RS}: response strength, X_R: response, X_{RT}: response trace, S_D: discriminative stimulus, X_{ST}: discriminative stimulus trace, S_R: reinforcement, w_{SM}: short-term memory for response-reinforcement associations, w_{LM}: long-term memory for response-reinforcement associations, X_{SE}: short-term learning expectancy, X_{LE}: long-term learning expectancy, X_{BI}: behavioral inhibition, X_{BE}: behavioral excitation. (+) excitatory (fixed) connections, (-) inhibitory (fixed) connections.





Responses and discriminative stimuli are assumed to leave traces which have the same time constant for both events (for simplicity, I ignore reinforcer traces). Figure 2 shows the dynamics of the model in a simplified situation in which a 3-sec response partially overlaps with the 2-sec reinforcement (discriminative stimuli are not considered, and therefore the superscripts, R and S in Figure 1B, indicating responses or stimuli, are missing). The onset of reinforcement triggers the formation of response-reinforcement associations (or response associative strengths), w^R_{SM} and w^R_{LM}, and stimulus-reinforcement associations (or stimulus associative strengths), w^S_{SM} and w^S_{LM}. These associations encode the temporal correlation between the trace of each response (X_{RT}) or stimulus (X_{ST}) and the reinforcement (S_R), i.e., the associative strengths are higher when the time lag between X_R or S_D and S_R is shorter. Response-reinforcement and stimulus-reinforcement associations are formed in parallel at two different time scales as short-term (w^R_{SM}, w^S_{SM}) and long-term memory traces (w^R_{LM}, w^S_{LM}) (Figures 1 and 2). When reinforcement ceases to occur both memory traces decay, with the short-term memory decaying faster than the long-term memory ($k_1 > k_2$).

One key assumption of the present theory is that reinforcement acts through a "novelty-detection" mechanism. *Novelty* means that the current experienced state of reinforcement must be compared with some "expected" state. The response and stimulus traces read out the short and long-term associations (which act as multiplicative "connection weights") to generate short-term ($X_{SE} = w^R_{SM}X_{RT} + w^S_{SM}X_{ST}$) and long-term ($X_{LE} = w^R_{LM}X_{RT} + w^S_{LM}X_{ST}$) *learning expectancies*. Notice that X_{SE} (short-term expectancy) is a measure of the experienced reinforcement and X_{LE} (long-term expectancy) is a measure of the expected reinforcement. These expectancies are generated only when a response is emitted or when a stimulus is "ON", and they gradually decay after the response ceases or the stimulus is turned "OFF".

Figure 2 *Typical behavior of model's various compartments.* Responses are generated continuously between time units 2 and 4. The reinforcement is presented between time units 4 and 5. X_R: response, X_{RT}: response trace, S_R: reinforcement, w_{SM}: short-term memory for response-reinforcement associations, w_{LM}: long-term memory for response-reinforcement associations, k_1 and k_2 are proportionality constants, X_{SE}: short-term learning expectancy, X_{LE}: long-term learning expectancy, X_{BI}: behavioral inhibition, X_{BE}: behavioral excitation, X_{RS}: response strength.

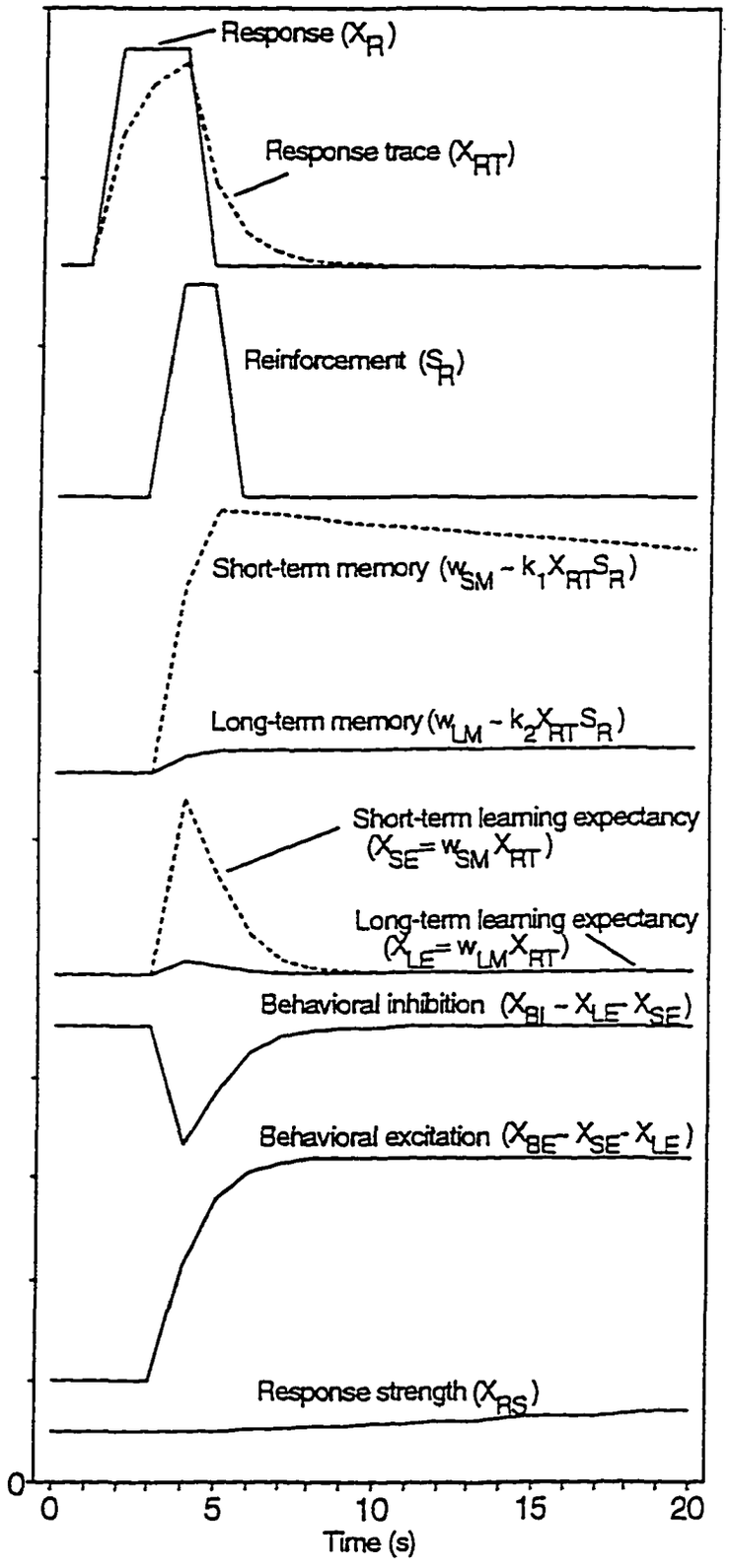


Figure 2 shows that X_{LE} and X_{SE} decay in the absence of responding. The mismatch between the short and the long-term learning expectancies controls the operant response: if the reinforcement is underpredicted (experienced reinforcement is larger than expected reinforcement, i.e., $X_{SE} > X_{LE}$) a *behavioral excitation* signal proportional to the difference between the short-term and the long-term expectancies ($X_{BE} \sim X_{SE} - X_{LE}$) enhances the response strength (X_{RS}); if the reinforcement is overpredicted (experienced reinforcement is smaller than expected reinforcement, i.e., $X_{SE} < X_{LE}$) a *behavioral inhibition* signal proportional to the difference between the long-term and the short-term expectancies ($X_{BI} \sim X_{LE} - X_{SE}$) reduces the response strength. Figure 3 shows that when the response and the reinforcement are active simultaneously (both short and long-term associations increase in strength) behavioral inhibition decreases and behavioral excitation increases ($X_{SE} > X_{LE}$). As soon as the reinforcement drops to 0 behavioral inhibition increases at a high rate whereas behavioral excitation reduces its rate of increase. Behavioral excitation will decrease as soon as the slowly decaying X_{LE} becomes greater than the rapidly decaying X_{SE} (in the current example the level from which the two expectancies decay prevents the two curves to cross). Finally, the end result of pairing the response and the reinforcement is the fact that the response gradually increases in strength (X_{RS} in Figure 2).

Theoretical principles

Response competition

Most theories of operant conditioning that deal with choice assume the existence of separate representations for each response, e.g., Luce (1959), Davis et al. (1993), Schmajuk (in press). I also represent alternate responses by separate response-strength units: X^i_{RS} in Figure 1. The level of the X^i_{RS} unit represents in real time the momentary preference for alternative i, where this alternative can be pecking key i, pressing lever i, etc. X_{RS} is functionally equivalent to the idea of response

strength or V-value used by other theories (Luce, 1959; Rescorla & Wagner, 1972; Staddon & Zhang, 1991; Mazur, 1992; Davis et al., 1993; Killeen, 1994).

In operant conditioning, every response followed by a reinforcer increases in strength (becomes more likely to recur). Consequently, the probability of responding to other alternatives (when they exist) must decrease. In other words, if X^i_{RS} increases then X^j_{RS} should decrease, and vice-versa. A simple implementation of this response-competition rule is the process of lateral (mutual) inhibition (see Figure 1). In the model, interresponse inhibition is mediated by fixed connections between the output units, X^i_{RS} and X^j_{RS} (see Equation 1, where the term $-\alpha_4 X^i_{RS} \sum_{j \neq i} X^j_{RS}$, represents response inhibition).

$$\frac{dX^i_{RS}}{dt} = -\alpha_1 X^i_{RS} + \alpha_2 X^i_{BE}(1 - X^i_{RS}) - \alpha_3 X^i_{BI} X^i_{RS} - \alpha_4 X^i_{RS} \sum_{j \neq i} X^j_{RS} \qquad (1)$$

X^i_{BE} is the output of the behavioral excitation unit, X^i_{BI} is the output of the behavioral inhibition unit, X^j_{RS} is the output of the "j"th competing response strength unit; α_1 controls the spontaneous decay of X^i_{RS}, α_2 controls the strength of excitation from X^i_{BE}, α_3 controls the strength of inhibition from X^i_{BI}, and α_4 controls the strength of mutual inhibition between responses.

The mutual inhibition process is behaviorally motivated by positive and negative contrast effects (Crespi, 1952; Reynolds, 1961; Gutman, 1977; Schwartz & Gamzu, 1977), which suggest that the extinction/facilitation of one alternative (or component of a multiple schedule) disinhibits/inhibits the response pattern to the other alternative (or component of a multiple schedule). The idea of response competition is generally accepted; there is disagreement only on exactly how it should be represented in a model. For instance, Herrnstein (1970) views the rate of response in one component of a multiple schedule as depending on the rate of reinforcement in the adjacent components, supporting thus a process of comparison (competition) of the value of one component to its neighbors. Staddon & Hinson (1978) use behavioral competition as a mechanism for schedule interaction. Davis et al. (1993) use a winner-take-all rule to model response selection, also supporting

the idea of nonlinear response competition.

The next step in defining the operant response is to convert the set of values of the X_{RS} units into discrete responses: to get from a configuration of X_{RS} values to an individual response. In conformity with the stochastic assumption used by other models of reinforcement learning, e.g., Luce (1959), Mazur (1992; 1995), it is assumed that an individual response X_R^i is generated with the probability $p(R_i) = X_{RS}^i / \sum_j X_{RS}^j$, where the sum from all X_{RS} units is taken. X_R^i is set to 1 if the subject responds to alternative i (according to the value of $p(R_i)$), and is 0 otherwise.

Short-term memory (STM) response trace

The existence of an STM response trace is justified by experimental results showing that operant conditioning can occur even if reinforcement is delayed after the response (e.g., Chung & Herrnstein, 1967; Killeen, 1968, 1970; McEwen, 1972): some effect of each response must persist for a short time. I assume that trace strength is directly related to response intensity and duration. (In the simulations, I assume that each response has a fixed unit intensity and duration). It is hypothesized that the trace of each response, X_{RT}^i, increases over time to a maximum during the response, and then decays to zero afterwards (see Equation 2; cf. Hull, 1943; Sutton & Barto, 1981, 1990).

$$\frac{dX_{RT}^i}{dt} = \alpha_5(X_R^i - X_{RT}^i) \tag{2}$$

where X_R^i and X_{RT}^i represent the response and the response trace, and α_5 is the rate of increase and decay of X_{RT}^i (see Appendix).

A similar equation can be written for the stimulus trace, X_{S_D}, where stimulus S_D replaces response X_R. In addition to responses and discriminative stimuli, I have also tested the idea of STM traces for reinforcement, motivated by evidence that reinforcers can be used as trace discriminative stimuli. However, since I obtained qualitatively similar simulation results with and without reinforcement STM

traces, their use did not seem to constitute a necessary condition and therefore I decided to keep the model as simple as possible.

Memory for response-reinforcement and stimulus-reinforcement associations

Following many suggestions from both experimental and computational neuropsychology (e.g., Grossberg, 1982; McClelland & Rumelhart, 1985; Schmajuk & DiCarlo, 1992) that simple stimuli leave brief traces whereas associations between stimuli (particularly associations between stimuli and reinforcement) leave more persistent traces, I assume that associations between response and reinforcement and stimulus and reinforcement leave both short, w^R_{SM} and w^S_{SM}, and long-term memory traces, w^R_{LM} and w^S_{LM}, where R stands for response associations and S stands for stimulus associations. The difference between short and long-term lies in the rate of change of the memory traces, i.e., w_{SM} varies faster than w_{LM}.

Short-term associations

Short-term memory for response-reinforcement and stimulus-reinforcement associations, w^R_{SM} and w^S_{SM}, is a measure of the currently experienced correlation between X_R and S_R and between S_D and S_R. It consists of a set of connection weights with small time constant which increase every time that a new association is formed, and then decrease until the occurrence of the following reinforced response (see Equation 3). If response X_R and stimulus S_D are followed by reinforcement, w^R_{SM} and w^S_{SM} increase and enhance the associations formed with the reinforcement. Equation 3 shows that the changes in $w^{R_i}_{SM}$ are driven by the product of X^i_{RT} and S_R, i.e.,

$$\frac{dw^{R_i}_{SM}}{dt} = -\alpha_6 w^{R_i}_{SM} + \alpha_6 X^i_{RT} S_R \tag{3}$$

where X^i_{RT} is the trace of X^i_R and α_6 controls the rate of increase and decay of $w^{R_i}_{SM}$ (see Appendix). A similar equation can be written for the STM for stimulus-

reinforcement associations, $w_{SM}^{S_i}$, where stimulus S_D replaces response X_R.

Long-term associations

Long-term memory for response-reinforcement and stimulus-reinforcement associations, w_{LM}^{R} and w_{LM}^{S}, is a measure of expected reinforcement. Similar to the short-term memory for associations, it consists of a set of connection weights which increase every time that a new association is formed, and then decrease until the occurrence of the following reinforced response. Equation 4 shows that the changes in $w_{LM}^{R_i}$ are driven by the multiplication between X_{RT}^{i} and S_R, i.e.,

$$\frac{dw_{LM}^{R_i}}{dt} = -\alpha_7 w_{LM}^{R_i} + \alpha_7 X_{RT}^{i} S_R \tag{4}$$

where α_7 controls the rate of increase and decay of $w_{LM}^{R_i}$, with $\alpha_7 << \alpha_6$ (see Appendix). A similar equation can be written for the LTM for stimulus-reinforcement associations, $w_{SM}^{S_i}$, where stimulus S_D replaces response X_R.

The similarity between w_{SM} and w_{LM} is that both reflect the strength of the associations between responses and reinforcement and between stimuli and reinforcement. The difference between these two parallel memory units is the time course of their integration: short-term memory integrates events over a small time scale, whereas long-term memory integrates events over a more extended time scale.

Learning expectancy

The response and stimulus traces, multiplied by their associative strengths, read out the corresponding short and long-term memory for associations and determine the short-term learning expectancy (a measure of experienced reinforcement) and the long-term learning expectancy (a measure of expected reinforcement). In Figure 1 the output nodes, X_{SE} and X_{LE}, represent the aggregate short and long-term learning expectancy at the level of one response unit.

I assume that the initial values of the short and long-term memory for response-reinforcement and stimulus-reinforcement associations are nonzero, i.e., w^R and w^S, see Equation 5. In this way, if a response or a stimulus occurs in the absence of reinforcement, w^R and w^S ensure that it can still have an effect (although small) on the overall learning expectancy, such that learning occurs at a lower rate. Evidence for this hypothesis is provided by the *latent learning* experiment (Tolman & Honzik, 1930) in which simply running rats through a maze with no food in the goal box does not prevent learning of a new response; rats not receiving any reinforcement can still learn to reach the goal box, but their performance (number of errors) is poorer than that of rats which were reinforced throughout the experiment.

Equation 5 expresses X_{SE_i} as the algebraic sum of all the STM modules at the level of each response, i.e.,

$$X^i_{SE} = X^i_{RT}(w^{R_i} + w^{R_i}_{SM}) + \sum_{j=1}^{N} X^j_{ST}(w^{S_j} + w^{S_j}_{SM}) \tag{5}$$

where $w^{R_i}_{SM}$ is the STM trace of X^i_R, $w^{S_j}_{SM}$ is the STM trace of the discriminative stimulus S^j_D, X^i_{RT} is the trace of X^i_R, X^j_{ST} is the trace of S^j_D, w^{R_i} is the initial (fixed) level of the connection between X^i_{RT} and X^i_{SE}, and w^{S_j} is the initial (fixed) level of the connection between X^j_{ST} and X^i_{SE}.

Equation 6 expresses X^i_{LE} as the algebraic sum of all the LTM modules at the level of response i, i.e.,

$$X^i_{LE} = X^i_{RT}(w^{R_i} + w^{R_i}_{LM}) + \sum_{j=1}^{N} X^j_{ST}(w^{S_j} + w^{S_j}_{LM}) \tag{6}$$

where $w^{R_i}_{LM}$ is the LTM trace of X^i_R, $w^{S_j}_{LM}$ is the LTM trace of the discriminative stimulus S^j_D, X^i_{RT} is the trace of X^i_R, X^j_{ST} is the trace of S^j_D, w^{R_i} is the initial (fixed) level of the connection between X^i_{RT} and X^i_{SE}, and w^{S_j} is the initial (fixed) level of the connection between X^j_{ST} and X^i_{SE}.

Behavioral Excitation and Behavioral Inhibition

Gray (1971, 1982), following Pavlov (1927), suggested that a behavioral inhibition system, activated by signals of punishment or nonreward, innate fear stimuli, or novel stimuli, generates an inhibitory signal which reduces the strength of the current response. In line with this approach, the model proposes an ensemble behavioral excitation - behavioral inhibition unit (X_{BE} and X_{BI} in Figure 1) whose main role is to detect variations in reinforcement contingency via short and long-term learning expectancies. If reinforcement conditions become better (larger size, smaller delay, increased duration, higher probability), short-term learning expectancy increases faster than long-term learning expectancy. If reinforcement conditions become worse (smaller size, larger delay, decreased duration, lower probability), short-term learning expectancy decreases faster than long-term learning expectancy. The X_{BE} unit integrates the difference (mismatch) between experienced (short-term) and expected (long-term) reinforcement, i.e., $X_{SE} - X_{LE}$, and the X_{BI} unit integrates the mismatch between expected (long-term) and experienced (short-term) reinforcement, i.e., $X_{LE} - X_{SE}$. In other words, X_{BE} increases whenever the reinforcement is underpredicted, i.e., $X_{SE} - X_{LE} > 0$, and decreases whenever the reinforcement is overpredicted, i.e., $X_{SE} - X_{LE} < 0$. X_{BI} increases whenever the reinforcement is overpredicted, i.e., $X_{LE} - X_{SE} > 0$, and decreases whenever the reinforcement is underpredicted, i.e., $X_{LE} - X_{SE} < 0$. After both X_{LE} and X_{SE} become 0, X_{BE} and X_{BI} slowly relax to 0. Equations 7 and 8 show that the mismatch between short and long-term reinforcement expectancies drives the dynamics of both X_{BE} and X_{BI},

$$\frac{dX^i_{BE}}{dt} = -\alpha_8 X^i_{BE} + \alpha_9 (X^i_{SE} - X^i_{LE})(1 - X^i_{BE}), \tag{7}$$

$$\frac{dX^i_{BI}}{dt} = -\alpha_{10} X^i_{BI} + \alpha_{11} (X^i_{LE} - X^i_{SE})(1 - X^i_{BI}), \tag{8}$$

where X^i_{LE} is the long-term learning expectancy, X^i_{SE} is the short-term learning expectancy, α_8 and α_{10} are rate constants of decay of X^i_{BE} and X^i_{BI}, and α_9 and

α_{11} are rate constants of increase of X^i_{BE} and X^i_{BI} (see Appendix).

Consolidation long-term memory

The term "consolidation long-term memory" is suggested by neuropsychological studies in which long-term memory is considered to have at least two components: labile long-term memory, which is recently acquired and easily disrupted by head injury as in retrograde amnesia (Russell, 1971), and fully consolidated long-term memory (Squire et al., 1975) which is more persistent and resistant to injury. I use consolidation long-term memory to encode very persistent effects.

Consolidation long-term memory, w_{CLM} (Figure 3), is updated using a correlation rule (proportional to the activity level of both X_{BE} and X_{RS}), such that w_{CLM} varies very slowly in time (small rate constant). Equation 9 describes the dynamics of the consolidation long-term memory:

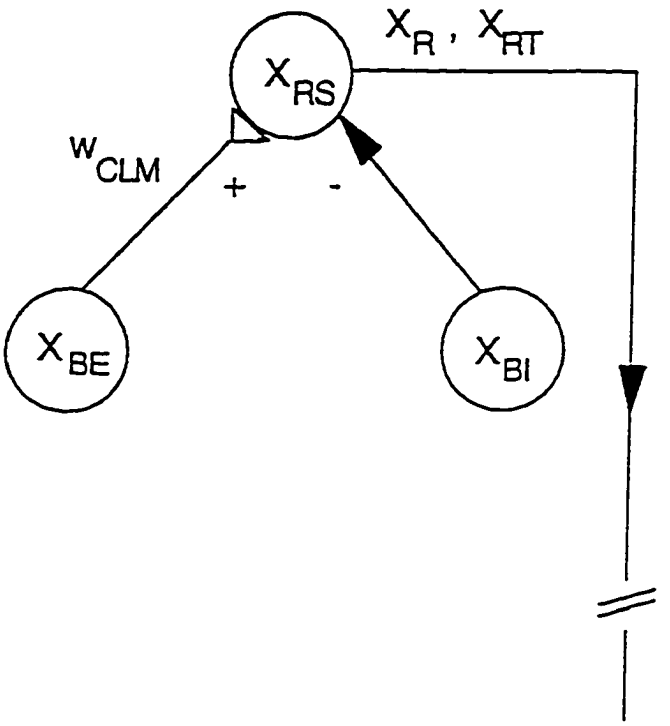
$$\frac{dw_{CLM}}{dt} = -\alpha_{12} w_{CLM} + \alpha_{13} X_{BE} X_{RS} \tag{9}$$

where α_{12} controls the rate of decay of w_{CLM} and α_{13} controls the rate of increase of w_{CLM}, with $\alpha_{12} < \alpha_{13}$ (see Appendix). X_{BE} is the behavioral excitation and X_{RS} is the response strength. As Equation 9 shows, consolidation long-term memory expresses the association between behavioral excitation and the response, and therefore it represents the cumulative effect of all stimuli associated with the reinforcement, including the response.

If $w_{CLM} \neq 0$ the strength of the "i"th response (X^i_{RS}) is given by the following equation:

$$\frac{dX^i_{RS}}{dt} = -\alpha_1 X^i_{RS} + (w_1 + w_{CLM}) X^i_{BE}(1 - X^i_{RS}) - \alpha_3 X^i_{BI} X^i_{RS} - \alpha_4 X^i_{RS} \sum_{j \neq i} X^j_{RS} \tag{10}$$

Figure 3 *Consolidation long-term memory.* Variable connection strengths between behavioral excitation and response strength (w_{CLM}).



where X^i_{BE} is the output of the behavioral excitation unit, X^i_{BI} is the output of the behavioral inhibition unit, X^j_{RS} is the output of the "j"th competing response strength unit; α_1 controls the spontaneous decay of X^i_{RS}, w_1 is the initial (fixed) level of the connection between X^i_{BE} and X^i_{RS} (numerically $w_1 = \alpha_2$ from Equation 1). The rest of the variables and parameters are identical to those from Equation 1.

Model dynamics

To illustrate the function of the model's various components, I present simulation results from (i) acquisition of partial reinforcement (Figure 4), and (ii) extinction (Figure 5). In these simulations the duration of each response and reinforcement is equal to 1 sec. Figure 4A shows a response sequence generated during a fixed-ratio 3 schedule (FR 3: three responses are needed in order to receive one unit of reinforcement). The partially reinforced sequence is: 10101100001111101001 1 ("1" = "response", "0" = "no response"), and runs for 20 time units (reinforcers occur at the black arrowheads). Each emitted response (X_R) activates the response trace, i.e., the X_{RT} representation. Figures 4B and 4C show the characteristic response trace dynamics (running average of emitted responses) and the obtained reinforcement (S_R) as a function of time, i.e., the S_R (reinforcement) sequence 10000100000010001000. Each reinforcement enhances the short and long-term memory for associations, or response associative strengths, (w_{SM} and w_{LM}) via a correlation rule established between the response trace and the reinforcement. Because the associative strengths reflect the contiguity between the response trace and the reinforcement (and not between the short-lasting response and the reinforcement), it can be used as a basis for delayed conditioning. Figure 4, panels D and E, shows the long and the short-term memory traces for response-reinforcement associations. According to the present theory, these memory traces differ in only one important respect: the rate of change of w_{SM} is high (small time constant) whereas the rate of change of w_{LM} is low (large time constant). This behavior

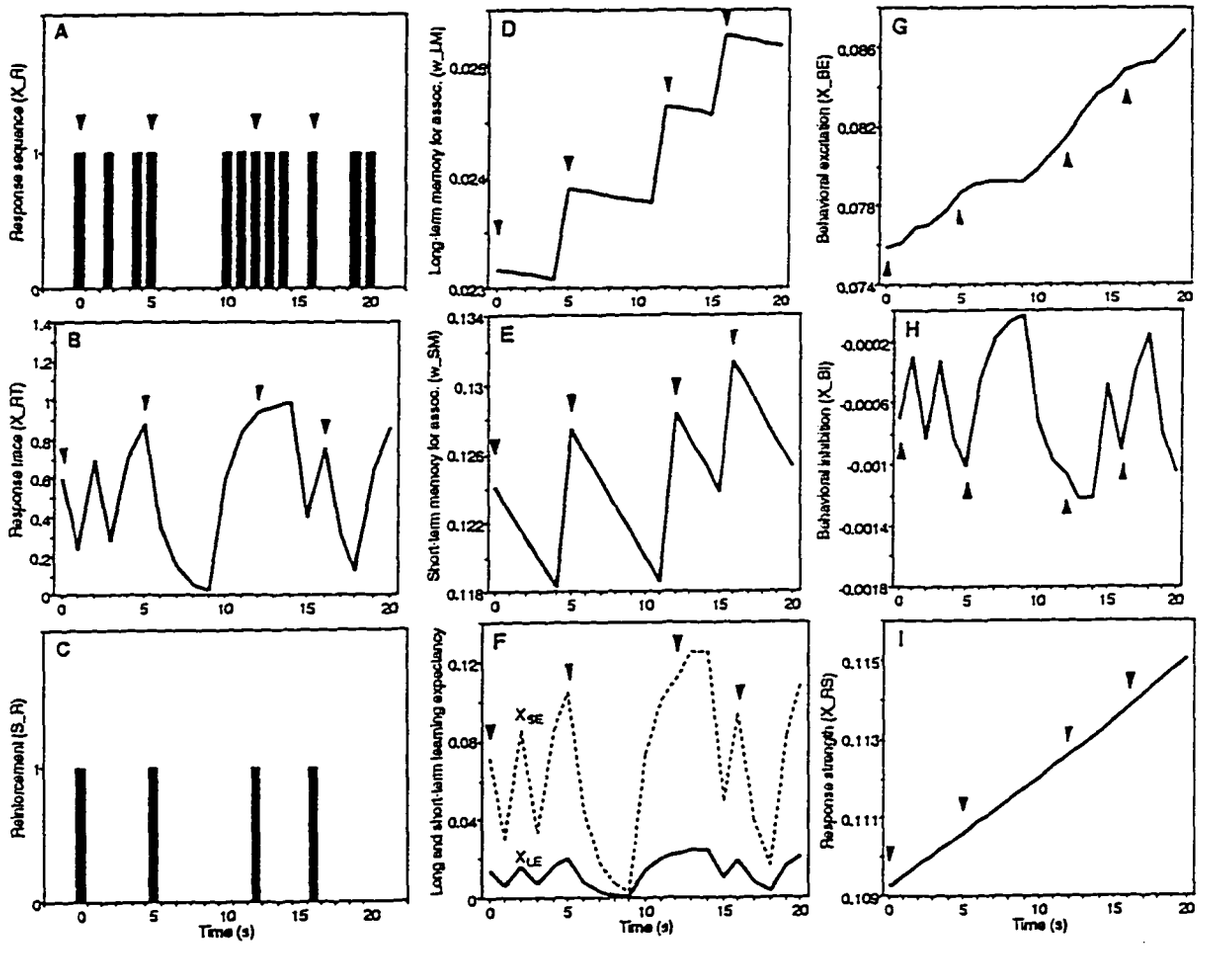
is clearly visible in Figure 4 (panels D-E): w_{LM} is perturbed to a lesser extent compared to w_{SM}, before and after each presentation of the reinforcement. Moreover, the high rate constant of the short-term memory trace allows it to vary at a higher rate compared to the long-term memory trace, a fact reflected in its larger magnitude. The magnitude of both short and long-term memory traces reflects associative strength and is influenced by response frequency, reinforcement frequency, reinforcement intensity and reinforcement duration.

The activity induced by the response gated by its short and long-term memory trace for associations determines the short-term learning expectancy (X_{SE}) and the long-term learning expectancy (X_{LE}). Figure 4F shows the dynamics of both X_{LE} and X_{SE}. Notice that the dynamics of both X_{LE} and X_{SE} are driven by the dynamic profile of the response trace (this is because the response reads out the memory traces to determine the level of reinforcement expectancy). The magnitude of change in both X_{LE} and X_{SE} is proportional to the activity level of long and short-term memory traces for associations (low for X_{LE} and high for X_{SE}).

The occurrence of variations in reinforcement contingency is detected by the comparison between expected (long-term) events and experienced (short-term) events. If the organism underpredicts the reinforcement (experienced events are better than expected) then both X_{LE} and X_{SE} increase, a situation which favors the increase in behavioral excitation (Figure 4G). If the organism overpredicts the reinforcement (experienced events are worse than expected) then both X_{LE} and X_{SE} start decaying. Since the decay process happens at different rates (high for STM traces and low for LTM traces), after a certain number of nonreinforced responses the more persistent memory for expected reinforcement becomes more salient than the less persistent memory for experienced reinforcement.

In this way, the change in the current reinforcement situation is detected by the behavioral inhibition unit (X_{BI}) which is driven by the difference between X_{LE} and X_{SE} (Figure 4H).

Figure 4 *Illustration of model dynamics during acquisition (FR 3 reinforcement schedule).* (A) Response sequence, X_R, generated during 20 time units (each reinforced response is marked with a black arrowhead). (B) Response trace, X_{RT}, dynamics. (c) Obtained reinforcement, S_R, as a function of time. (D) Long-term memory trace for response-reinforcement associations, w_{LM}. (E) Short-term memory trace for response-reinforcement associations, w_{SM}. (F) Dynamics of long and short-term learning expectancy, X_{LE} and X_{SE}. (G) Behavioral excitation, X_{BE}, increases whenever the short-term learning expectancy is greater than the long-term learning expectancy (for instance during training) and decreases otherwise. (H) Behavioral inhibition, X_{BI}, increases whenever the long-term learning expectancy is greater than the short-term learning expectancy (for instance during extinction) and decreases otherwise. (I) Response strength, X_{RS}, increases with the accumulation of reinforcement.

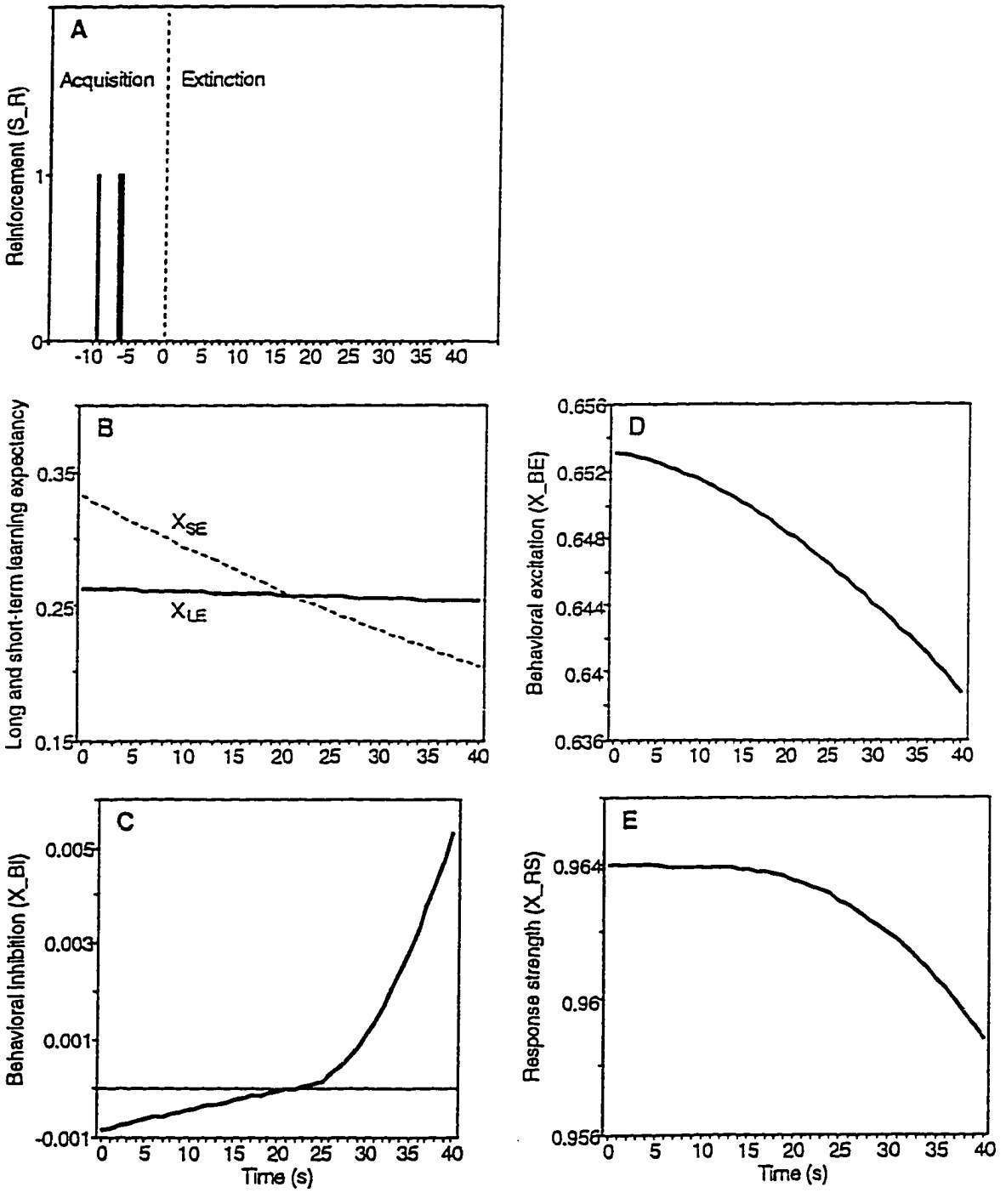


The output of the behavioral inhibition unit increases whenever both long-term and short-term expectancies decrease (for instance during extinction), and decreases otherwise. It is easy to observe that X_{BI}'s profile shown in Figure 4H is out of phase with respect to the temporal dynamics of both X_{LE} and X_{SE}. As mentioned previously, the rate of change of the expectancy units is such that short-term expectancy varies faster than long-term expectancy, i.e., $|\frac{dX_{LE}}{dt}| < |\frac{dX_{SE}}{dt}|$. When both expectancy units increase (the first time derivatives are positive) then $\frac{dX_{LE}}{dt} < \frac{dX_{SE}}{dt}$, and therefore $\frac{d(X_{LE}-X_{SE})}{dt} < 0$, i.e., the expectancy mismatch decreases, a fact that contributes to the decrease in the amplitude of behavioral inhibition, see Figure 4 (panels F-H). When both expectancy units decrease (the first time derivatives are negative) then $-\frac{dX_{LE}}{dt} < -\frac{dX_{SE}}{dt}$, and therefore $\frac{d(X_{LE}-X_{SE})}{dt} > 0$, i.e., the expectancy mismatch increases, a fact that favors the increase in behavioral inhibition (see Figure 4, panels F-H).

Figure 4I shows that when behavioral inhibition is negligible, like in the current simulation, behavioral excitation controls the rate of change of the response strength unit, X_{RS} (the response strength increases with the accumulation of reinforcement).

Figure 5 shows the behavior of representative model elements during extinction. Figure 5A shows that the last reinforcement is presented at time -6 (the FR 3 schedule is in effect until time 0, the moment when extinction starts). Figure 5B illustrates the decay of both short and long-term learning expectancy units, with X_{LE} decaying slower than X_{SE}. As soon as $X_{LE} > X_{SE}$, behavioral inhibition increases to positive values (Figure 5C) causing the decrease in response strength that matches the extinction situation. During extinction behavioral excitation also decreases (Figure 5D) due to the decay of both short and long-term expectancies. The net result of these effects is the response strength which gradually decreases (Figure 5E).

Figure 5 *Illustration of model dynamics during extinction.* (A) The FR 3 schedule is in effect until time 0, the moment when extinction starts (no reinforcement delivered thereafter). (B) Long and short-term learning expectancies decrease at different rates: slow (X_{LE}) and fast (X_{SE}). (c) During extinction, behavioral inhibition, X_{BI}, becomes positive and suppresses the response strength. (D) Behavioral excitation decreases during extinction due to the decay of learning expectancy. (E) Response strength, X_{RS}, decreases due to the suppressive effect of behavioral inhibition.



The Dynamics of Operant Conditioning

The present model is intended to account for the major phenomena of operant learning, particularly in recurrent choice situations with food reinforcement. The data set is the real-time pattern of responses under different manipulations of reinforcement probability as a function of time. I attempt to explain qualitative, and when possible quantitative, patterns of change in real-time operant responding. The theory allows us to model both steady-state and transient response during acquisition and extinction under various conditions. In many situations predictions of the model are compared with experimental data. However, I will show only predictions without comparing them with data when the effects Iaim to explain reflect a well-known general property of operant conditioning which has been observed in many experimental conditions and in different species or when the lack of complete procedural and performance information (e.g., other than "trials to criterion") makes detailed comparison impossible.

One difficulty which I experienced throughout simulations was to accommodate (using the same set of parameters) all the quantitative aspects of the data referenced in this article, given that the data ought to be explained emerges from experiments with various time-courses, ranging from hours, even days (e.g., spontaneous recovery) to one session (e.g., development of preference), and from experiments involving various species and responses (e.g., pigeons, rats, or starlings and pecking, perch-hopping or alley-running). I therefore show quantitative predictions only for free-operant pigeon experiments with few training sessions. But in other cases - involving complex, incompletely specified or long-time-course procedures – I had to settle for qualitative matches between data and predictions. Nevertheless, I have done sufficient computer simulations (not presented in this article) to make sure that it is possible to match most real time courses by careful selection of an appropriate integration time step.

All the simulations in this section involve one or two operant responses, a reinforcement, and a discriminative stimulus for each response. Unless otherwise men-

tioned, all the simulation results were obtained using fixed connection strengths between behavioral excitation and response strength. When the computer simulations start, the activity level of all the variables is initialized to the same small random value (this manipulation is needed to eliminate any bias in responding). The reinforcement conditions are signaled by the discriminative stimulus which switches from 0 to 1 and triggers one response at random.

The Appendix shows the numerical values of the parameters, held constant for all simulations. Each unit in the model has a fixed discharge rate (in the absence of any stimulation) equal to 0.0001. Loosely speaking, I have implemented three different time-scales: short (response trace, interresponse inhibition strength, short-term memory for associations), long (long-term memory for associations), and very long (consolidation long-term memory). Correspondingly, I have limited the parameter search to time courses of the following orders: 10^{-2} (short), 10^{-3} (long), and $10^{-4} - 10^{-5}$ (very long). With these constraints, it was relatively easy to find a configuration of values that generates curves whose profile and relative time course approximately fit the profile of experimental data. The parameter set indicated in the Appendix probably represents one of the best configurations with respect to graphical appearance. However, the dynamics of the model are quite robust to perturbations in the parameter space. I found that the most sensitive parameters are those controlling the response strength unit, i.e., the connections between $X_{BI} - X_{RS}$, $X_{LE} - X_{RS}$, and $X_{RS} - X_{RS}$.

Response Rule, Reinforcement Contingency

Each next response, X_R^i, is generated with the probability $p(R_i) = \frac{X_{RS}^i}{\sum_{j=1}^{N} X_{RS}^j}$ from the set of N available responses. If response X_R^i is generated, a random number between 0 and 1 is compared with the reinforcement probability for response i. If the random number is smaller or equal to the reinforcement probability, a reinforcement is set to 1 for one time unit (the same method used in experimental conditions during a concurrent VR-VR schedule). In the case of concurrent VI-VI

schedules, the reinforcement is set up independently for each response alternative, and, once a reinforcement is set up, it remains available until collected. In the case of discrete-trial experiments, each emitted response, either reinforced or not, is followed by a fixed ITI (intertrial interval) during which the discriminative stimulus is turned OFF. After the ITI, the next trial begins by turning ON the discriminative stimulus until the first response is recorded.

Assignment of Credit

When reinforcement is contingent on a particular response, the probability of that response generally increases; when reinforcement ceases, becomes less frequent or is presented independent of responding, response probability should decrease. These issues are incorporated under the *Assignment of Credit* problem, a first step toward the investigation of the mechanism through which the reinforcement selectively strengthens the "reinforced" response, a question which any theory of operant conditioning should attempt to answer. I will explain the major aspects of the Assignment of Credit problem based on the simulation results.

Response selection

Experimental data. Response selection is the first prerequisite of a dynamic model for operant conditioning. Sutton & Barto's (1990) temporal-differences model and Staddon & Zhang's (1991) assignment-of-credit model address response selection. However, most other operant conditioning models (e.g., Myerson & Miezin's, 1980, kinetic model, Killeen's, 1994, model) take response selection for granted and do not show explicitly how reinforcement selects a response by decreasing the strengths of alternative responses as a function of contingency.

Response selection can be analyzed in a simple two-armed bandit situation. In this situation (concurrent VR-VR schedule) responses on a rich and a lean side are paid off with different probabilities (higher for the rich side). The problem can be stated as follows: (a) How does the organism select the "rich" response without

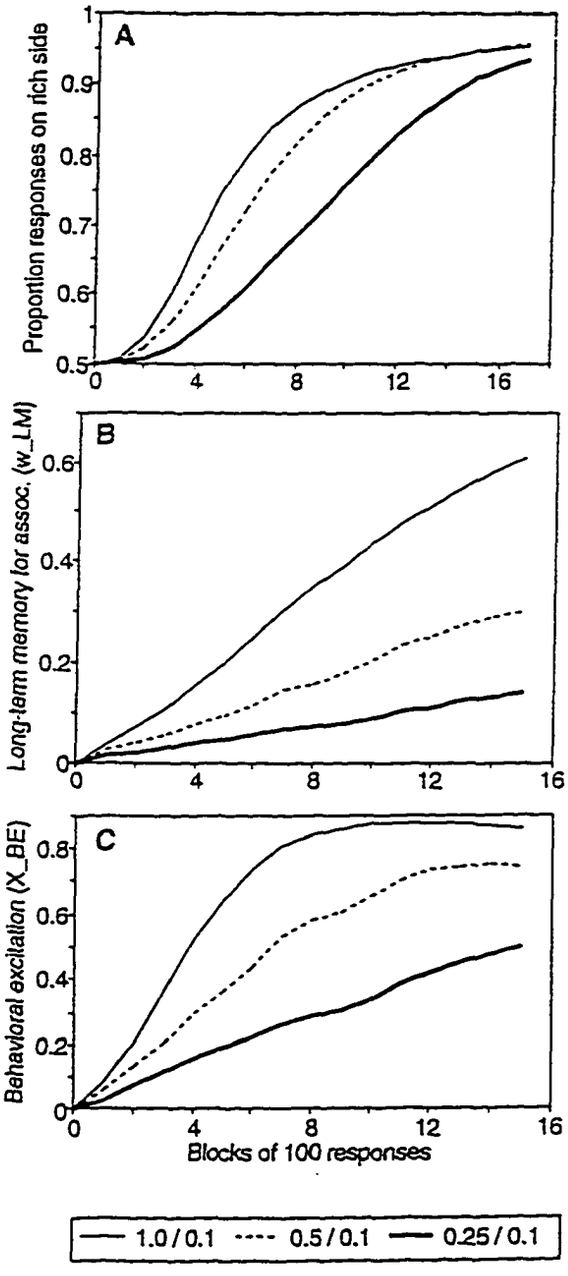
explicit a priori "knowledge" of what that response is? (b) Why is the "rich" response selected faster when reinforcement probability on the rich side increases?

Simulation results. To answer these questions, I consider a generic situation in which three different conditions are run under a concurrent VR-VR schedule. The lean side provides reinforcement with fixed probability, 0.1 (VR 10), while the rich side is set at three different levels, i.e., 1.0 (CRF), 0.5 (VR 2), and 0.25 (VR 4). Figure 6A shows the proportion of responses on the rich side in each of these conditions as a function of blocks of 100 responses. As the reinforcement probability of the rich side increases, the rate of response selection also increases. This result is explained as follows. Let us assume that initially the subject samples both alternatives with equal probability, i.e., both X_{RS} units are set to the same level.

Figure 6B shows the profile of the long-term memory trace for response-reinforcement associations[3] for the "rich" side, from which it can be concluded that responses reinforced at a higher rate tend to form stronger associations with reinforcement, w_{LM} (cf. Equation 4). Figure 6C shows that behavioral excitation (X_{BE}) for the rich side increases at a rate controlled by the payoff probability. This result is explained by the fact that the increase of both short and long-term memory traces for associations determines the increase of the learning expectancy units (X_{SE} and X_{LE}) at a rate proportional to the reinforcement probability. Therefore, behavioral excitation, which integrates the difference between the short and the long-term learning expectancy units, facilitates response strength unit by controlling its rate of increase. In this way, the "rich" response is selected faster with the increase in learning expectancy.

[3]Short-term memory trace profile is similar except for the higher rate constant.

Figure 6 *Response selection.* (A) Proportion of responses on the rich side as a function of blocks of 100 responses. Responses reinforced at a higher rate are selected faster (choice approaches fixation faster at higher reinforcement rates). (B) Responses reinforced at higher rates develop stronger associations with the reinforcement, w_{LM}. (c) Rate of increase of behavioral excitation is positively correlated with the rate of increase of long-term memory for response-reinforcement associations.



At the same time, the competition between operant responses ensures that the strength of responses on the lean side wanes at the same time that responses on the rich side become stronger. In this way, after a sufficient number of responses the subject fixates on the rich side at a rate controlled by the corresponding reinforcement probability on the rich side (Figure 6A).

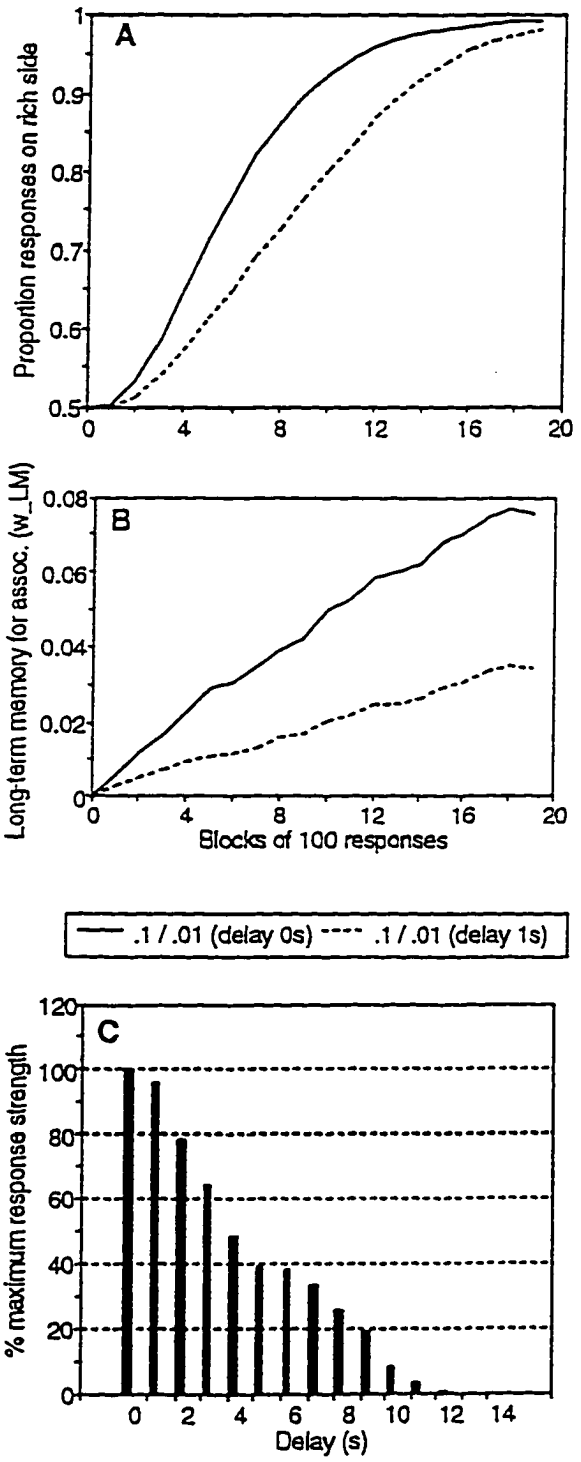
Delayed reinforcement

Experimental data. Both operant and classical conditioning data show that if the reinforcement (or unconditioned stimulus) is delayed, performance is usually disrupted (e.g., Logan, 1960; Renner, 1963). In both discrete-trial and free-operant situations, animals show a preference for the alternative associated with a shorter delay of reinforcement (Chung & Herrnstein, 1967; Killeen, 1968, 1970; McEwen, 1972) even though the probability of reinforcement on the two alternatives is the same.

Simulation results. Figure 7A presents a typical situation in which the effects of reinforcement delay are analyzed in a concurrent VR 10 - VR 100 schedule in which reinforcement for responses on the rich side is delayed (1 sec). I observe that the proportion of responses on the rich side in a 0 delay situation increases faster compared to the situation in which the rich side is reinforced with a 1-s delay.

According to the theory, the reinforcement strengthens short and long-term associations. The strength of response-reinforcement associations (Eqns. 3 and 4) depends on the magnitude of the product $S_R \cdot X_{RT}$, where S_R is the reinforcement and X_{RT} is the response trace. In the interval between the emission of the last response and the occurrence of the contingent reinforcement, the trace of response gradually decreases. Therefore, any delay in reinforcement is equivalent to the increase in phase between response and reinforcement, event that causes the formation of weaker response-reinforcement associations (Figure 7B) that generate a weaker learning expectancy.

Figure 7 *Delayed reinforcement.* Responses are reinforced with probability 0.1 (rich side) and 0.01 (lean side), respectively. (A) The proportion of responses on the rich side in a 0-s delay situation increases faster compared to the proportion of responses in a 1-s delay situation. (B) Reinforcement delay causes the formation of weaker long-term response-reinforcement associations. (C) Delay-of-reinforcement-gradient in which the percentage of maximum response strength decreases with the reinforcement delay. Reinforcement probabilities are the same as above (0.1/0.01).



In these conditions the delayed-reinforcement response receives diminished support from the behavioral excitation unit, thus generating a lower level of responding. Figure 7C presents a delay-of-reinforcement-gradient form of the effect presented in panel A.

Preference Reversal and Delayed Reinforcement

Experimental data. In the previous subsection I analyzed how preference develops for the shorter delay of reinforcement when both reinforcement magnitude and probability are held constant across conditions. In an attempt to test the properties of the hyperbolic discount curve proposed to describe how the reinforcement loses its value per unit of delay time, Ainslie & Herrnstein (1981) analyzed the pattern of preference when both reinforcement delay and magnitude vary. They showed that in a two-key discrete-trial procedure with the larger reinforcement always presented 4 sec later than the smaller, as the delay between the access to the smaller reinforcement is varied from .01 to 12 sec, pigeons reverse preference from the small-early to the large-late reinforcement (Figure 8A).

Simulation results. I have simulated the discrete-trial procedure used by Ainslie & Herrnstein, running one set of computer simulations for each pair of early-late reinforcement delays such that the ratio between late and early reinforcement magnitudes is two. I explain the preference reversal effect as a result of the interplay between the different time scales used in the model. Intuitively, preference for each side is determined by the product between the response trace (X_{RT}) and the size of the reinforcement (S_R) associated with each response. The response trace gradually decreases with the delay between the response and the reinforcement. Since the decay of the response trace is a negatively accelerated function (cf. Equation 2), when the reinforcement delay is small the associations between the early-small reinforcement and the (high level) response trace are stronger than those between the late-large reinforcement and the (low level) response trace. However, if the reinforcement delay increases the absolute difference between response traces in

the small-early vs. large-late situation diminishes, and therefore the associations between the early-small reinforcement and the response become weaker than those between the large-late reinforcement and the response, thus accounting for the shift in preference with the increase in reinforcement delay (Figure 8A-B).

Note that the model is able to explain delayed-reinforcement preference reversal without assuming the hyperbolic discount curve proposed for this effect (Ainslie, 1975; Chung & Herrnstein, 1967). I explain the effect as an emergent property of the interaction between processes that occur at different time scales.

Effect of Free Reinforcement

Experimental data. Contingency means that the strengthening effect of reinforcement is a function of the degree of correlation with the reinforced behavior, not just an effect of contiguity. In addition to the effects explained in previous subsections, an important instance of contingency is the decrease in response rate with the delivery of free reinforcements. The effect of free reinforcement has been little addressed in formal operant theorizing (but see Staddon & Zhang, 1991).

Simulation results. I have simulated three variable-ratio simple schedules, VR 60, VR 80 and VR 120; as soon as stable response rates were obtained free reinforcers are delivered on three different variable-time schedules (VT 10-s, VT 40-s, and VT 160-s) while the maintaining VR schedule was in effect. Figure 9 shows that because the free reinforcers degrade the response-reinforcement contingency the rate of responding decreases with respect to baseline across the three conditions. It can be also shown that, consistent with Cohen et al. (1993) findings, the decline in the rate of responding is positively correlated with the probability of free reinforcement. There are greater reductions with VT 10-s schedules than with VT 40-s or VT 160-s schedules.

Figure 8 *Preference reversal and delayed reinforcement.* (A) The pattern of preference when both reinforcement delay and reinforcement magnitude vary. In two-key discrete-trial procedures, with the larger reinforcement always presented 4 sec later than the smaller, as the delay between the access to the smaller reinforcement is varied from .01 to 12 s, pigeons reverse preference from the small-early to the large-late reinforcement. [adapted from Ainslie & Herrnstein (1981)]. (B) Simulation results obtained in the same conditions as in the original experiment.

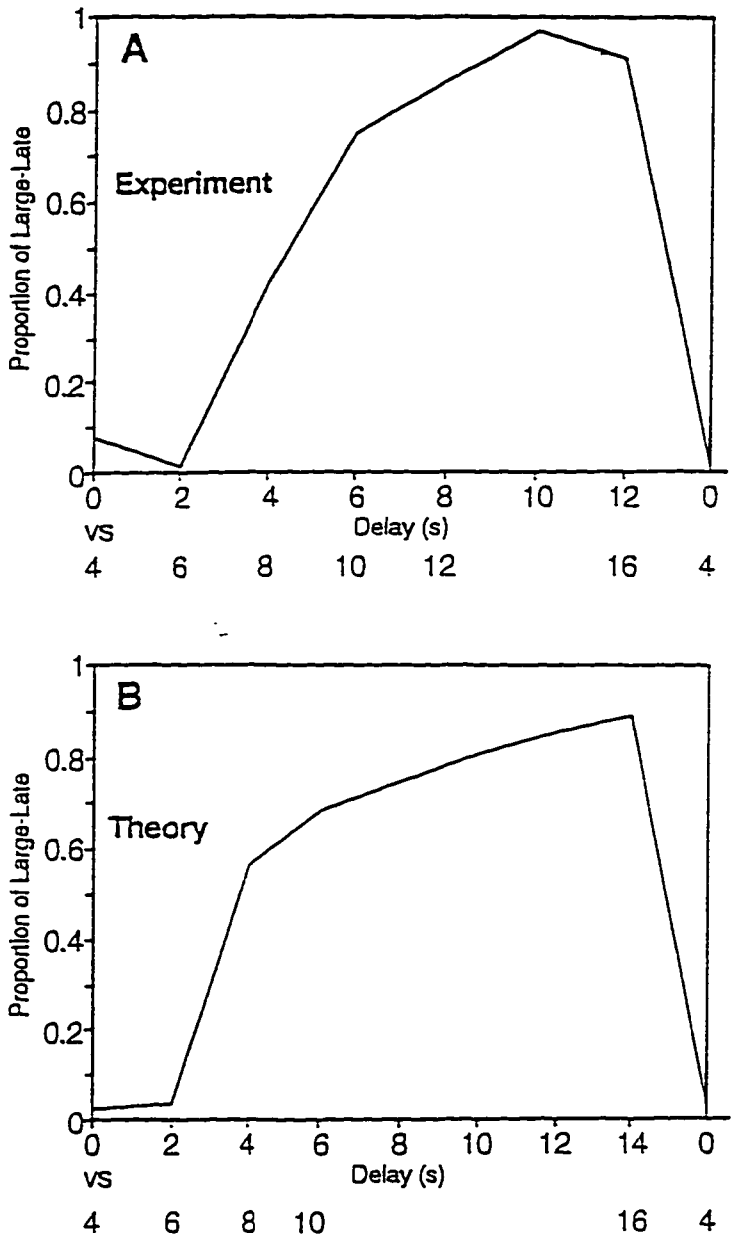
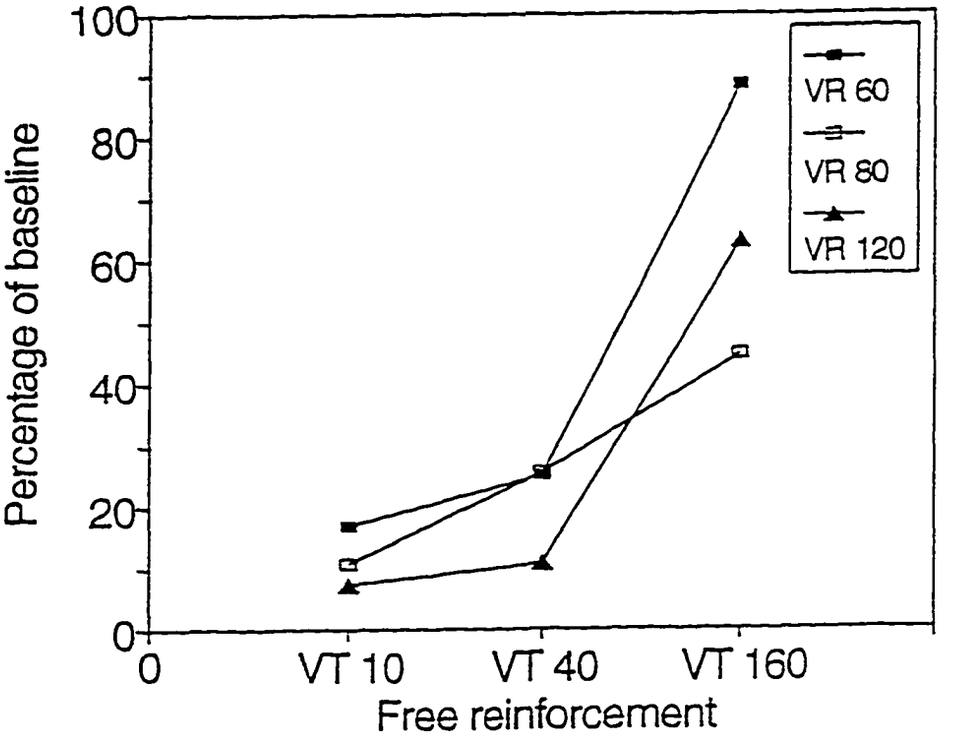


Figure 9 *Effect of free reinforcement.* The percentage of response rate during a free-reinforcement test to rate on the immediately preceding baseline session (percentage of baseline). Data is obtained in three different conditions (VR 60, VR 80, VR 120) during training. In the test phase free reinforcement is delivered, while the maintaining schedule is in effect, according to the following schedules: VT 10-s, VT 40-s, or VT 160-s. The decline in response rate is positively correlated with the probability of free reinforcement.

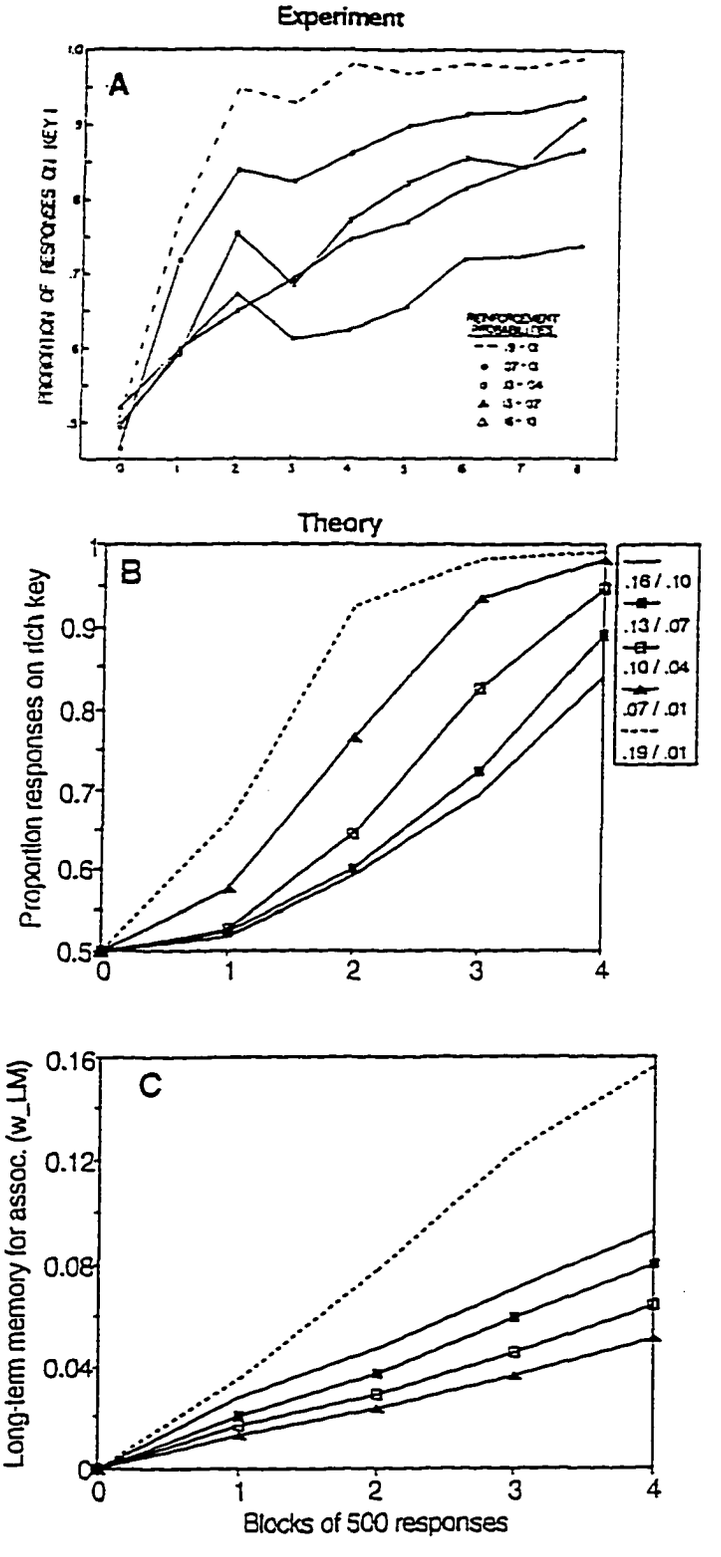


To explain the free-reinforcement effect, I consider that the explicitly reinforced response competes throughout training with responses representing the set of other behaviors from the animal's innate repertoire (cf. Staddon & Simmelhag, 1971). I group and label these responses as o-response or "other" response (I ensure that throughout simulations the probability of the o-response does not fall behind a 0.03 threshold level). During the standard training conditions, the response-reinforcement contingency ensures that as the reinforcement accumulates the contingent response increases in strength. However, if no response is emitted the trace of the contingent response decays until it reaches 0. If a reinforcement is offered in these conditions it becomes associated with, and strengthens, the o-response. As free reinforcers continue to accumulate, the expectancy of o-response - reinforcement associations increases and causes the strength of s-responses to grow at a rate controlled by the probability of free reinforcement. Because of response competition, the o-response inhibits further the contingent response that decreases in strength.

Development of Preference

In simple choice situations such as concurrent variable-ratio schedules (VR-VR) choice shifts towards the side with the higher reinforcement probability. Choice data obtained in both discrete-trial experiments, e.g., Davenport (1962), and free-operant experiments, e.g., Kacelnik et al. (1987), support the following generalization: the higher the probability (or magnitude) of reinforcement for the rich side the faster the convergence toward the rich alternative, a result that seems completely intuitive. How does the distribution of the two reinforcement probabilities affect the *rate* of transition toward the "winning" alternative? Specifically, what is the effect of (a) The ratio between reinforcement probabilities (the two probabilities vary such that their ratio is held constant)? (b) The absolute difference between reinforcement probabilities (the two probabilities vary such that their absolute difference is held constant)?

Figure 10 *Development of preference when the reinforcement probability ratio is varied but the absolute difference is constant.* (A) Proportion of responses on the rich key as a function of blocks of 500 responses. Even though the difference between the two probabilities is .06 for each of the four probability pairs, the transition to preference for the higher probability of reinforcement develops faster with the larger ratio between the two reinforcement probabilities [adapted from Mazur & Ratti (1991) - Figure 1]. (B) Simulation results obtained in conditions similar to the original experiment. Preference increases faster with the larger ratio between the two reinforcement probabilities. (c) Long-term memory for associations between responses to the rich key and the reinforcement. Even though lower probability ratios are equivalent to stronger associations, however preference is an increasing function of probability ratio.



Effect of ratio between reinforcement probabilities

Experimental data. Mazur & Ratti (1991) suggested that the development of preference for one alternative depends on the discriminability of the two alternatives, showing that with concurrent VR-VR schedules the acquisition of preference occurs more rapidly with larger ratios between the probabilities of reinforcement, even when their absolute difference is held constant. For instance, suppose that in one condition the two reinforcement probabilities are .16 and .10, and in another condition the two probabilities are .07 and .01. Even though the difference between the two probabilities is .06 in both conditions (Figure 10A), the transition to preference for the higher probability of reinforcement is much faster in the .07/.01 condition (ratio 7) than in the .16/.10 condition (ratio 1.6).

Simulation results. I have simulated the transition sessions of the free-operant choice experiment described by Mazur & Ratti (1991). Each response on one key (rich) was reinforced with a probability p_1, and each response to the other key (lean) was reinforced with a probability $p_2 < p_1$. By keeping $p_1 - p_2$ fixed (equal to 0.06), while varying p_1/p_2, the development of preference is compared across four groups. The pairs of reinforcement probabilities are: 0.16/0.10, 0.13/0.07, 0.10/0.04, and 0.07/0.01. The fifth pair, 0.19/0.01 was presented to give subjects an easy discrimination that can be used for control purposes. Figure 10B shows simulation results that demonstrate that preference develops faster with the larger ratios between the two reinforcement probabilities. Figure 10C shows that, for the situations described in the current simulation, even though lower reinforcement probability ratios are equivalent to higher response-reinforcement long-term memory associations, however preference is an increasing function of p_1/p_2.

This effect can be explained as a conjoint result of response competition and behavioral excitation. According to the theory, during acquisition the rate of increase of X_{RS_1} ('1' is the richer side) is influenced by two factors (cf. Equation 1). One is facilitatory, the term $\alpha_2 X^1_{BE}(1 - X_{RS_1})$ which sustains acquisition, and one is suppressive, the term $-\alpha_4 X^1_{RS} X^2_{RS}$ which determines the strength of competi-

tion between the two responses (suppressive influences from behavioral inhibition are negligible during acquisition). It was shown earlier that behavioral excitation (X_{BE}) is proportional to learning expectancy and that learning expectancy is proportional to the frequency of contingent reinforcement. Therefore, the higher the ratio between the two reinforcement probabilities (p_1/p_2) the higher the ratio between the corresponding X_{BE} units (X^1_{BE}/X^2_{BE}) which sustain the two competing responses. Since behavioral excitation controls the rate of increase of the response, higher ratios between the X_{BE} units (corresponding to the rich and the lean sides) will cause higher ratios between the level of the corresponding X_{RS} units (X^1_{RS}/X^2_{RS}). Notice that this dependency holds only in transition, not in the steady state of the X_{RS} units, as Equation 1, which describes X_{RS}, is a saturating equation with respect to the term which contains X_{BE}. According to the response rule, a singular response, X^1_R, is emitted with probability $p(R_1) = X^1_{RS}/(X^1_{RS} + X^2_{RS})$, a function that varies monotonically with X^1_{RS}/X^2_{RS}. If the ratio p_1/p_2 increases, then X_{RS_1}/X_{RS_2} also increases, yielding a higher probability, p_1, for response X^1_R, i.e., more frequent responses on the rich side and less frequent responses on the lean side. Therefore, there will be less inhibition on X^1_{RS} by X^2_{RS}, resulting in a faster preference for the richer side (preference for alternative 1 is defined as $X^1_{RS}/(X^1_{RS} + X^2_{RS})$).

Effect of absolute difference between reinforcement probabilities

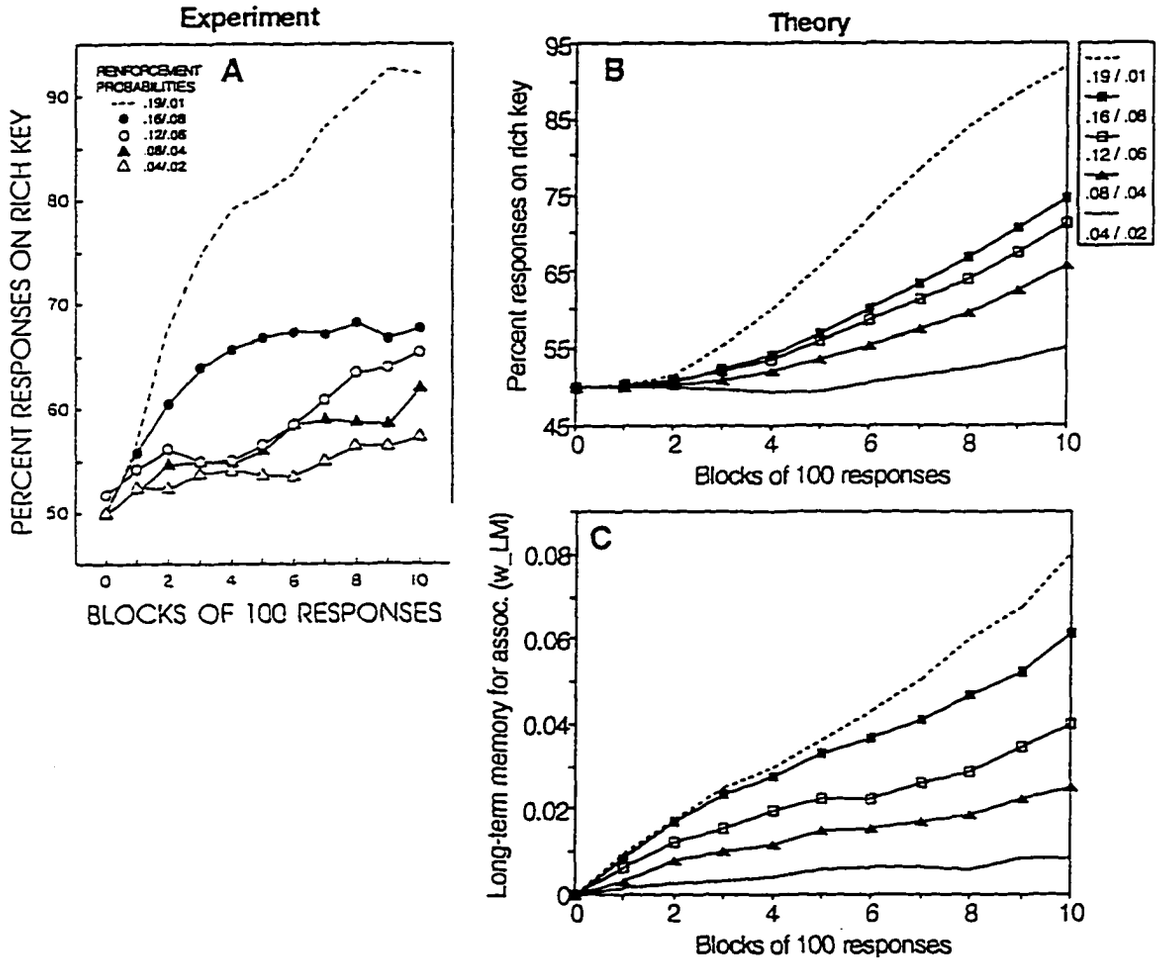
Experimental data. Mazur (1992) showed that when the ratio between the two probabilities of reinforcement is held constant, preference develops according to the absolute difference between the two probabilities. For instance, suppose that in one condition the two reinforcement probabilities are .16 and .08, and in another condition the two probabilities are .08 and .04. Even though the ratio between the two probabilities is 2 in both conditions, the transition to preference for the higher probability of reinforcement is much faster in the .16/.08 condition (difference .08) than in the .08/.04 condition (difference .04). Mazur interprets these findings

by assuming a Weber-law type of effect in the acquisition of preference, i.e., two stimuli (reinforcement probability) are more easily discriminated (or processed) if they differ by a larger percentage or absolute difference.

Simulation results. I have simulated the transition sessions of the discrete-trial choice experiment described by Mazur (1992). Each response on one key (rich) was reinforced with a probability p_1, and each response to the other key (lean) was reinforced with a probability $p_2 < p_1$. Keeping p_1/p_2 fixed (equal to 2), while varying $p_1 - p_2$, allows the development of preference to be compared across four groups. The pairs of reinforcement probabilities are: 0.16/0.08, 0.12/0.06, 0.08/0.04, and 0.04/0.02. The fifth pair, 0.19/0.01 was presented to give subjects an easy discrimination that can be used for control purposes. Figure 11B illustrates simulation results showing that preference develops faster with the larger absolute difference between the two reinforcement probabilities. Figure 11C shows that lower absolute differences between reinforcement probabilities are equivalent to lower reinforcement expectancies.

As in the case of probability ratios, the absolute-difference effect can be explained as a conjoint result of response competition and behavioral excitation. Given that the learning expectancy varies proportionally to reinforcement probability, at equal ratios between reinforcement probabilities one should expect equal ratios between learning expectancies. However, if the absolute difference between reinforcement probabilities increases (at the same ratio), the probability of the rich side also increases. This leads to the formation of stronger response-reinforcement associations for the richer alternative, causing the level of X^1_{BE} to increase in time to a higher level. This contributes to the facilitation of X^1_{RS} which increases faster and sends more inhibition to X^2_{RS}, thus sustaining a higher rate of fixation. Mazur's (1992) results can be explained by means of the excitatory effect of the absolute reinforcement probability for the rich alternative.

Figure 11 *Development of preference when the reinforcement-probability absolute difference is varied but the ratio is constant.* (A) Proportion of responses on the rich key as a function of blocks of 100 responses. Even though the ratio between the two probabilities is 2 for each of the four probability pairs, the transition to preference for the higher probability of reinforcement develops faster with the larger absolute difference between the two reinforcement probabilities [adapted from Mazur (1992) - Figure 2]. (B) Simulation results obtained in conditions similar to the original experiment. Preference increases faster with the larger absolute difference between the two reinforcement probabilities. (c) Long-term memory for associations between responses to the rich key and the reinforcement. Lower absolute differences between reinforcement probabilities are equivalent to weaker associations with the reinforcement.



Mazur's experimental results are inconsistent with most other models of acquisition, for example, the linear-operator model (Bush & Mosteller, 1955), the kinetic model (Myerson & Miezin, 1980), melioration theory (Herrnstein & Vaughan, 1980), ratio invariance (Staddon, 1988), etc., although Grossberg (1972), in his analysis of punishment and avoidance, did derive a Weber law in the development of the avoidance response. Other dynamic models such as the CE model or DMOD (Daly & Daly, 1982), fail to generate correct predictions when applied to this phenomenon. The theory shows that the critical variables to explain Mazur's results are response competition and the facilitatory effect of learning expectancy (expressed via behavioral excitation).

Matching

Experimental data. Herrnstein (1961) proposed a general principle, known as the *matching law*, which states that in a concurrent VI-VI schedule the proportion of responses for one alternative should equal the proportion of reinforcement delivered by that alternative. Matching is not perfect: The most common deviation is undermatching (Baum, 1979; Davison & McCarthy, 1988), in which response proportions are consistently less extreme than reinforcement proportions. Undermatching can be reduced experimentally by introducing a *changeover delay* (COD), which is a reinforcement-delay penalty for switching between alternatives (Herrnstein, 1961). Previous simulation results have shown that matching is overdetermined: almost any law-of-effect process (more reinforcement leads to more preference) will produce something like matching on concurrent variable-interval schedules (Hinson & Staddon, 1983), so it would be surprising if the model failed to predict matching. More interesting would be prediction of systematic deviations from perfect matching.

Simulation results. I simulated a concurrent VI-VI schedule in which the percentage of reinforcements on one side was varied between 0 to 100%. I analyzed two situations, COD=0 and COD=2 sec. Figure 12 presents experimental data in

parallel with theoretical results. Percentages of responses and reinforcements for one alternative were calculated over the last 10 simulation sessions (1 session = 60 reinforcements), after preference has stabilized. The typical finding is that undermatching is obtained in the absence of COD. However, Figure 12 shows that undermatching is reduced if the COD is set to 2 sec, a value close to that used by Herrnstein (1961) to obtain perfect matching. The results of the simulation are compared with Herrnstein's data (bird 055) in Figure 12, the match is almost point for point.

These results are mildly surprising because, like the cumulative-effects model (Davis et al., 1993), it shows that a nonlocal process can produce matching (most previous matching models are local, e.g., Herrnstein & Vaughan, 1980; Myerson & Miezin, 1980; Horner & Staddon, 1987). However, unlike the cumulative-effects model which is insensitive to real time and therefore is unable to deal with the COD procedure, the present model does show that a 2-sec COD reduces undermatching. Nonetheless, because the theory has no way to represent learning of switching (i.e., use of a response as discriminative stimulus), it cannot show that large increases in COD almost abolish switching.

Kinetics of matching

Experimental data. In an attempt to analyze the time course of matching, Mark and Gallistel (1994) used concurrent VI-VI schedules with brain-stimulation reward in rats to measure trial-to-trial changes in the relative rates and relative magnitudes of reward. The experimental conditions were repeated sessions of two trials each, the relative scheduled rate of reward reversing from 4:1 in favor of one lever to 4:1 in favor of the other between the two trials. The transition to matching was traced cumulating the times each lever was depressed and the number of rewards received within successive temporal windows equal in width to the expected interreward interval of the leaner schedule.

The results (Figure 13A) support Dreyfus' (1991) findings, i.e., window-to-

window variations in the time-allocation ratio track the random unsignaled window-to-window variations in the experienced reward rate, results independent of the time scale of the experiment. Moreover, the transition to matching is almost complete within the first time window of the post-transition trial. However, in a subsequent set of experiments, Mark & Gallistel (1995) replicated earlier results showing that if the animal experiences stable relative rates of reward for many sessions before the relative rate is reversed, the adjustment in time allocation ratio is slower and is less sensitive to the noise in the reinforcement schedule than if relative reward rates are held constant only briefly (Figure 13B). The conclusion of Mark & Gallistel's studies is that the development of matching behavior depends on past experience, with slower transition to matching being observed if the animal has experienced stable reward rates for extended sessions. (Note that the qualitative result here - faster preference changes when reward ratios change more frequently - is exactly the same as a familiar result in serial reversal learning that I discuss in a moment, namely faster reversal when contingencies change daily vs. every two or four days [Davis & Staddon, 1990]).

Simulation results. I have simulated Mark & Gallistel's (1994; 1995) experiments using a concurrent VI 4-s VI 16-s schedule. The relative scheduled reward rate reverses from 4:1 in favor of one alternative to 4:1 in favor of the other in the middle of each session. During each part of the session I have calculated the ratio between the number of rewards received on each side and the ratio between the number of responses on each side within time windows approximating two expected interreward intervals on the leaner side. Successive reward and response ratios are plotted on a logarithmic scale as a function of session time.

Figure 12 *Matching on concurrent variable-interval (VI) schedules.* Dashed line: theoretical matching law (the proportion of responses for one key should equal the proportion of reinforcement delivered by that alternative). Rectangles: Herrnstein data [taken from Herrnstein (1961)]. Thin line: simulation results show undermatching (response proportions are less extreme than reinforcement proportions). Heavy line: Corrections of undermatching can be obtained by introducing a *changeover delay* (COD), which is a penalty for switching between alternatives. In the simulations I used a 2-s COD. Percentages of responses and reinforcements for each alternative are calculated for the last 10 simulation sessions.

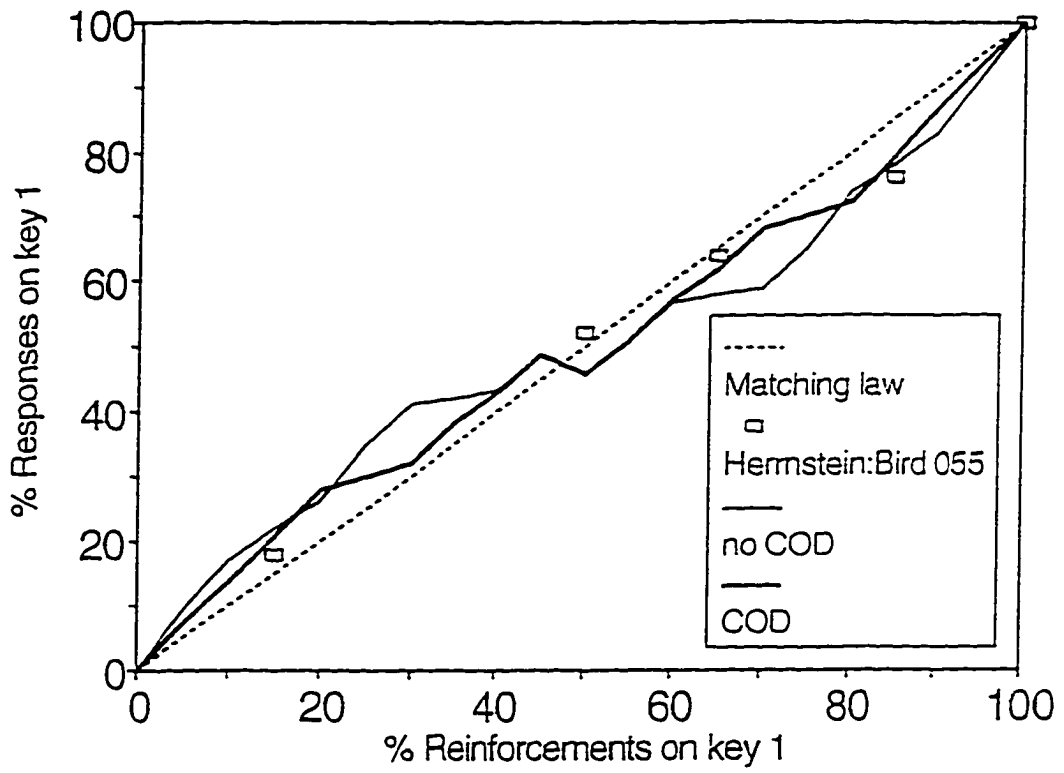
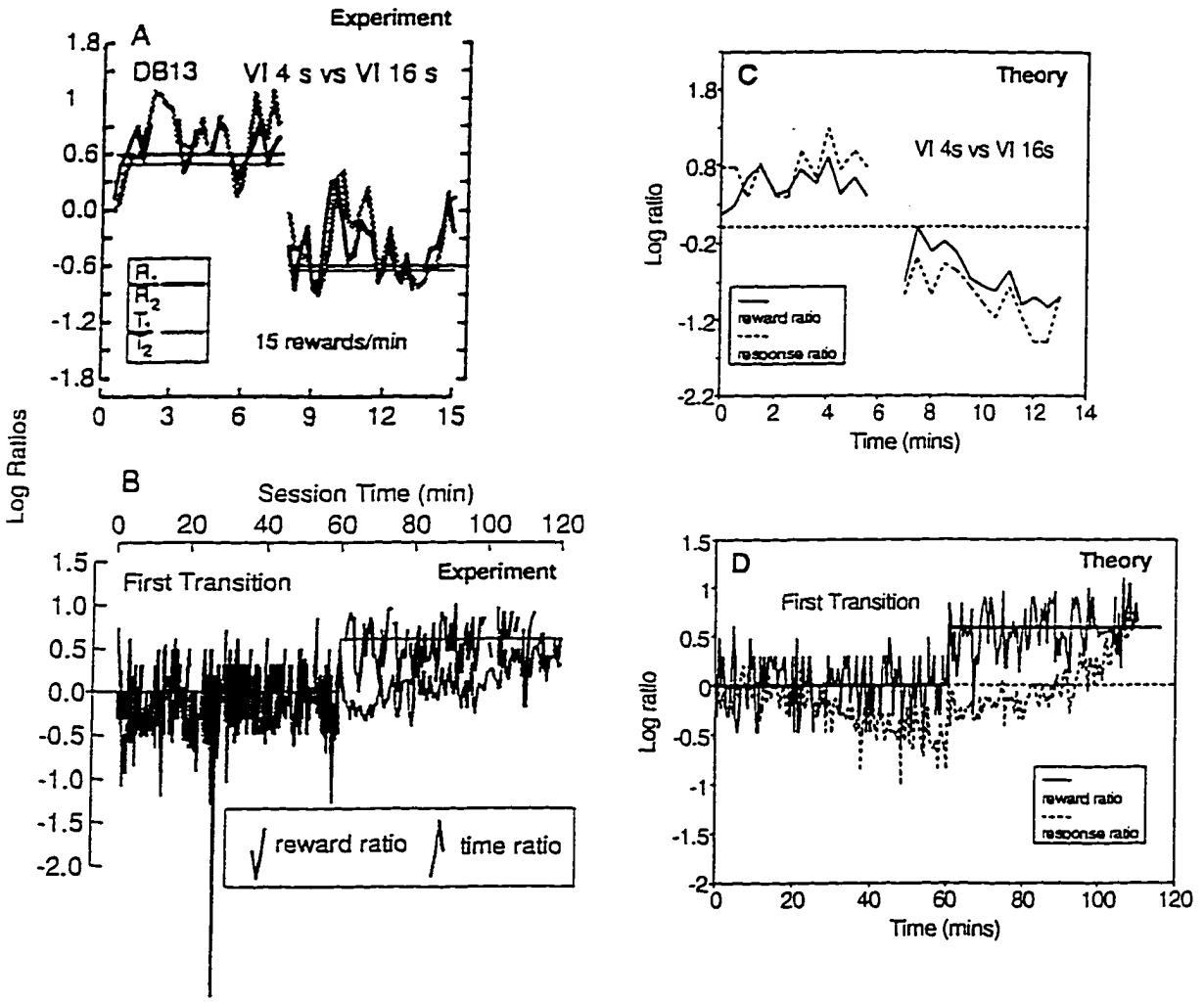


Figure 13 *Kinetics of matching.* (A) Rats were exposed to concurrent VI-VI schedules with brain stimulation reward. The experimental conditions were repeated sessions of two trials each, the relative scheduled rate of reward reversing from 4:1 in favor of one lever to 4:1 in favor of the other between the two trials. The transition to matching was traced cumulating the times each lever was depressed and the number of rewards received within successive temporal windows equal in width to the expected interreward interval of the leaner schedule. [adapted from Mark & Gallistel (1994) - Figure 1A] (B) If rats experience stable relative rates of reward for many sessions before the relative rate is reversed, the adjustment in time allocation ratio is slower and is less sensitive to the noise in the reinforcement schedule compared to the situation in which relative rates of reward are held constant only for small durations. [adapted from Mark & Gallistel (1995) - Figure 7] (C) Simulation results of Mark & Gallistel's (1994) experiment in a concurrent VI 4-s VI 16-s procedure. The transition to matching is almost complete within the first time window of the post-transition trial. (D) Simulation results of Mark & Gallistel's (1995) experiment in a concurrent VI 4-s VI 16-s procedure. Matching is slower if the animal has experienced stable reward rates for extended sessions.



The results are consistent with Mark & Gallistel's (1994) findings, i.e., the response ratios match the reward ratios a few minutes into each session. Furthermore, response ratios show fluctuations that track the noise in the reward ratios. This result can be explained as an effect of the short-term time scales in the model that offer the substrate for rapid matching: if the time window in which reward and response ratios are calculated is short, responses on each side reflect the effective obtained reward rates, and matching occurs as a normal consequence. However, if the change in the scheduled reward rate occurs after extended training (60 min in the simulation) the development of matching takes much longer, results which are consistent with Mark & Gallistel's (1995) findings. In this situation, the longer time scales in the model tend to preserve the consequences of the previous reinforcement conditions. Specifically, when the change in the scheduled reward rates occur, the long-term memory for response-reinforcement associations decays slowly in response to the new reinforcement conditions, overshadowing the effect of the rapid decline in the short-term memory for associations which becomes thus less effective.

Successive contrast effects

The term contrast effect refers to those situations in which exposure to one reinforcement condition affects performance under a succeeding condition. Contrast effects have been consistently demonstrated in both runway and discrete-trial free-operant studies with rats and pigeons: Crespi (1952), Nevin & Shettleworth (1966), Bernheim & Williams (1967), Franchina & Brown (1971).

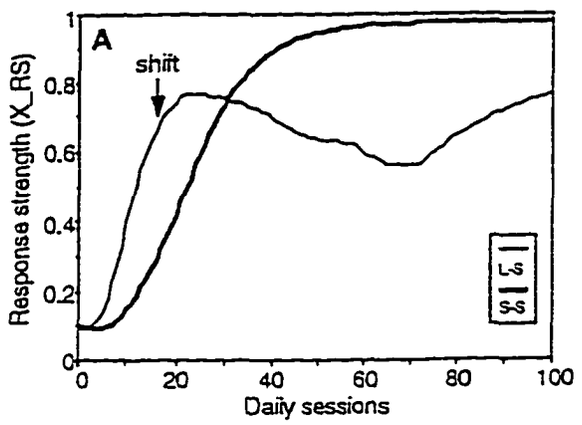
Successive negative contrast effect

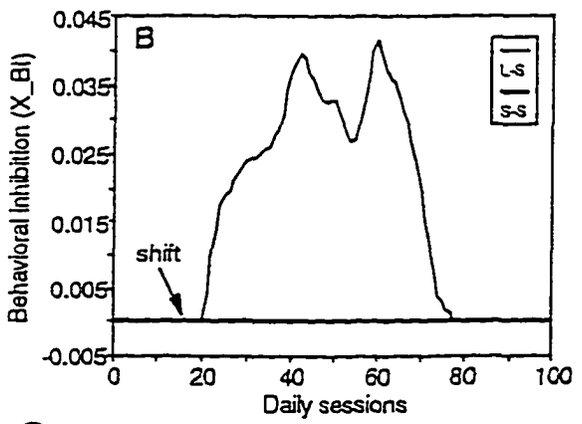
Experimental data. If subjects are switched from a CRF schedule (continuous reinforcement, i.e., every response reinforced) to a PRF schedule (partial reinforcement, i.e., the response is reinforced with some probability), PRF response rate is depressed (and approaches asymptote more slowly) compared to a control group

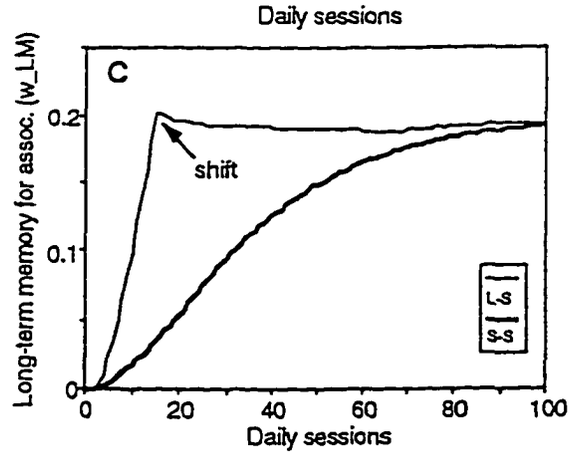
that received reinforcement with the same PRF probability throughout (successive negative contrast effect: Crespi, 1952; Black, 1968; Cox, 1975). Negative contrast is transient; the response rate gradually recovers from depression.

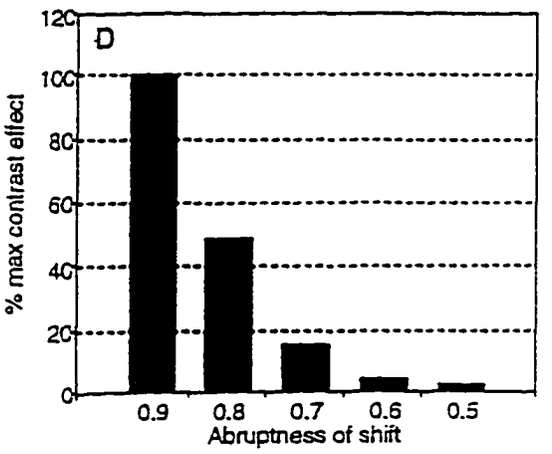
Simulation results. To test the basic effect, I simulate a discrete-trial procedure in which the model is exposed to 100 acquisition sessions, with 15 reinforcements per session. During each trial animals receive either a large (L) or a small (S) reinforcement (ratio large:small, 3 : 1). The control group ($S - S$) receives a small reinforcement for all the 100 sessions. The "shifted" group ($L - S$) receives a large reinforcement for the first 18 sessions, and then was switched to the small reinforcement for the rest of the sessions. Figure 14A shows that after the shift occurs, the performance (response strength) gradually becomes higher in the control group. According to the theory, when the negative shift occurs the animal overpredicts the reinforcement. The difference in the time constants at which the short and long-term memory traces for associations decay, as an effect of reinforcement diminution, contributes to a positive mismatch between the long and the short-term learning expectancy units ($X_{LE} - X_{SE} > 0$) that determines a positive behavioral inhibition (X_{BI}) signal (Figure 14B) that reduces the response strength for the "shifted" group. Figure 14C shows that despite the discontinuity that follows session 18, the "shifted" group has a higher level of response-reinforcement associations than the control group. However, after the mismatch between short and long-term memory for associations diminishes, behavioral inhibition gradually decreases and releases the response strength unit from depression. The simulation results (Figure 14A) show gradual changes in response strength following the shift in reinforcement magnitude. This result is similar to many negative contrast studies (e.g., Meyer, 1951; Spence, 1956; Bower, 1961; Di Lollo & Beez, 1966) reporting gradual rather than abrupt changes in performance. The gradual changes suggest that the development of the negative contrast effect is driven by learning processes (as hypothesized here), distinct from motivational variables.

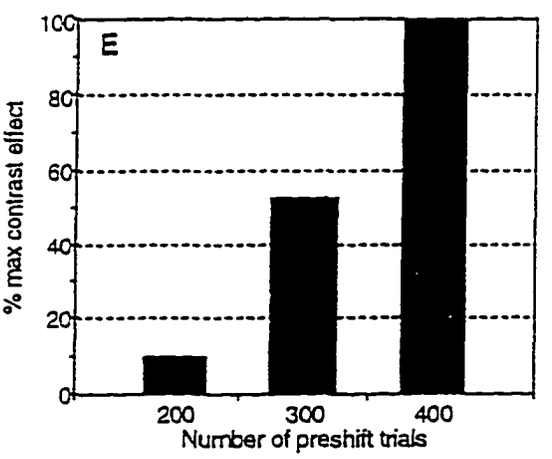
Figure 14 *Successive negative contrast effect.* (A) Response strength is calculated in two different situations: (1) small reinforcement for all sessions as control condition (S-S); (2) large reinforcement (L) for the first 18 sessions followed by small (S) reinforcement for the rest of sessions as "shifted" condition (L-S). Even though the response is much stronger for the L-S segment (compared to the S-S case), once the reinforcement magnitude diminishes the response strength of the "shifted" group develops at a lower rate than the response strength of the control group. (B) Behavioral inhibition detects the change in reinforcement magnitude and inhibits the response of the "shifted" group. (c) Long-term memory for response-reinforcement associations: the "shifted" group has a higher reinforcement expectancy than the control group, despite the abrupt change that follows trial 18. (D) Magnitude of effect: negative contrast effect is an increasing function of the abruptness of shift (difference between the large and the small reinforcement). (E) Magnitude of effect: negative contrast effect is an increasing function of the number of preshift trials.











The present theory gives an accurate picture of the effects on negative contrast of abruptness of shift, number of preshift trials, and quality and delay of reinforcement. It is shown here the influence of the difference between large and small reinforcement (abruptness of shift) and the influence of the number of trials preceding the shift. Figure 14D shows that as the abruptness of shift is varied between 0.9 and 0.5 the depression effect reduces in strength, a result consistent with Mikulka et al's (1967) findings, who showed that negative contrast effect is reduced if the reinforcement discontinuity is made less abrupt.

Figure 14E shows that the negative contrast effect is an increasing function of the number of preshift trials (Vogel et al., 1966). This result can be explained by the fact that extended preshift acquisition trials contribute to a larger reinforcement expectancy that, once the reinforcement magnitude is decreased, determines a larger behavioral inhibition signal that reduces the response strength.

Successive positive contrast effect

Experimental data. If subjects reinforced on a PRF schedule are switched to CRF they usually perform at a higher level (response rate) than a control group exposed only to CRF (elation effect or positive contrast effect: Benefield et al., 1974; Maxwell et al., 1976). The positive contrast effect is transient; the response rate gradually recovers from elation.

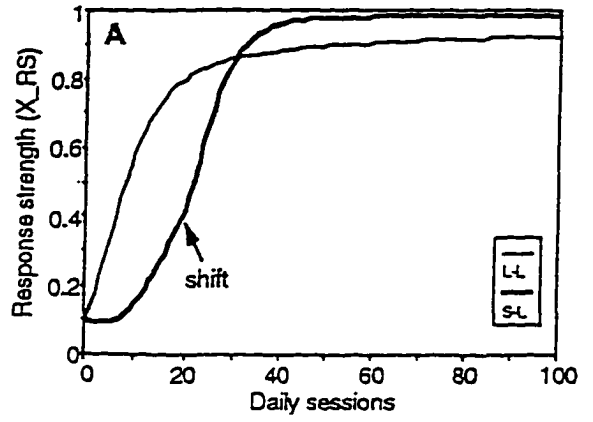
Simulation results. I simulate a discrete-trial procedure in which the model is exposed to 100 acquisition sessions, with 15 reinforcements per session. During each trial animals receive either a large (L) or a small (S) reinforcement (ratio large:small $= 3 : 1$). The control group ($L - L$) receives a large reinforcement for all the 100 sessions. The "shifted" group ($S - L$) receives a small reinforcement for the first 20 sessions, and then is switched to the large reinforcement for the rest of the sessions. Figure 15A shows that after the shift occurs performance (response strength) gradually becomes lower in the control group. This result is explained as a direct consequence of reinforcement underprediction, i.e., the sudden

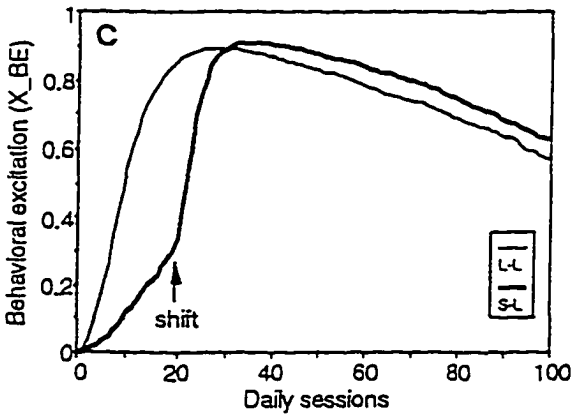
increase in reinforcement magnitude causes the short-term learning expectancy to increase faster than the long-term learning expectancy, causing further behavioral excitation (X_{BE}) to increase above the control level (Figure 15B) despite the fact that behavioral excitation starts from a lower level in the "shifted" condition. It is noteworthy that despite increased response strength in the "shifted" condition, stronger response-reinforcement associations are formed in the control condition (Figure 15C), where the large reinforcement is in effect since the first acquisition trial. This result shows once again that the properties of operant response are a function of changes in the level of response-reinforcement associations (reflected in the dynamics of behavioral excitation and behavioral inhibition) rather than a function of the absolute level of these associations. Notice that the strength of the operant response cannot be predicted solely from the asymptotic level of the short and long-term associations.

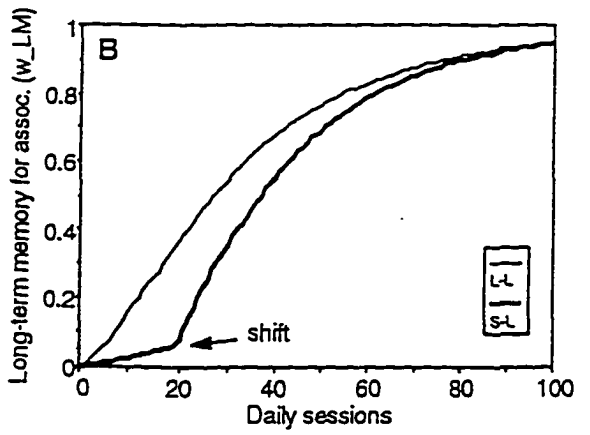
It is easy to see from simple inspection (Figures 14A and 15A) that for the same magnitude of the shift in reinforcement conditions the simulations show an asymmetry between negative and positive contrast effects. The magnitude of successive negative contrast effect is larger than the magnitude of positive contrast. However, although at the first sight this result might seem implausible, the available evidence supports these findings. Despite the fact that positive contrast was long thought to be symmetric and as readily obtainable as negative contrast, investigations that followed Crespi's (1952) initial report in successive experiments (e.g., Spence, 1956; Sgro & Weinstock, 1963; Capaldi & Lynch, 1967) either failed to find positive contrast or found only a relatively small effect that points to the conclusion that there is no positive contrast effect equal and opposite to the negative contrast effect generally observed when animals are shifted from a large to a small reinforcement.

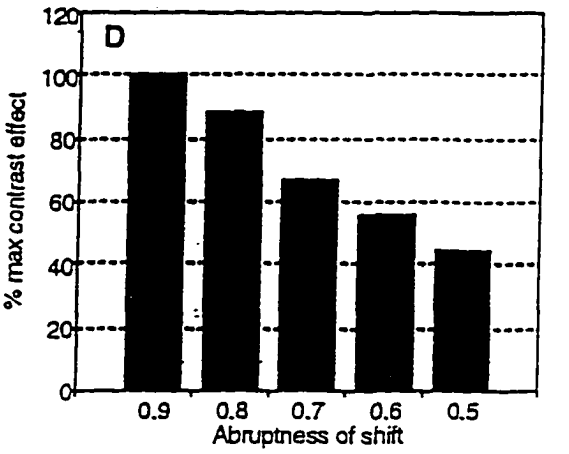
The theory correctly characterizes the determinants of positive contrast effects, i.e., abruptness of shift, reinforcement delay and frequency.

Figure 15 *Successive positive contrast effect.* (A) The response strength is calculated in two different situations: (1) large reinforcement for all sessions as control condition (L-L); (2) small reinforcement (S) for the first 20 sessions followed by large (L) reinforcement for the rest of sessions as "shifted" condition (S-L). Even though the response is much stronger for the L-L segment (compared to the S-L case), once the small reinforcement is shifted to a large one the response strength of the "shifted" group develops at a higher rate than the response strength of the control group. (B) Behavioral excitation detects the change in reinforcement magnitude and increases the amount of facilitation to the response of the "shifted" group. (c) Long-term memory for response-reinforcement associations: the "shifted" group has a lower reinforcement expectancy than the control group, despite the fact that the response strength develops to higher levels. (D) Magnitude of effect: positive contrast effect is an increasing function of the abruptness of shift (difference between the large and the small reinforcement).









I present here the influence of the absolute difference between the small preshift reinforcement and the large postshift reinforcement (abruptness of shift). Figure 15D shows that as the shift in reinforcement magnitude is varied between 0.9 and 0.5 the positive contrast effect declines, a result consistent with the findings of Weisman (1969) who showed that positive contrast is enhanced by larger shifts.

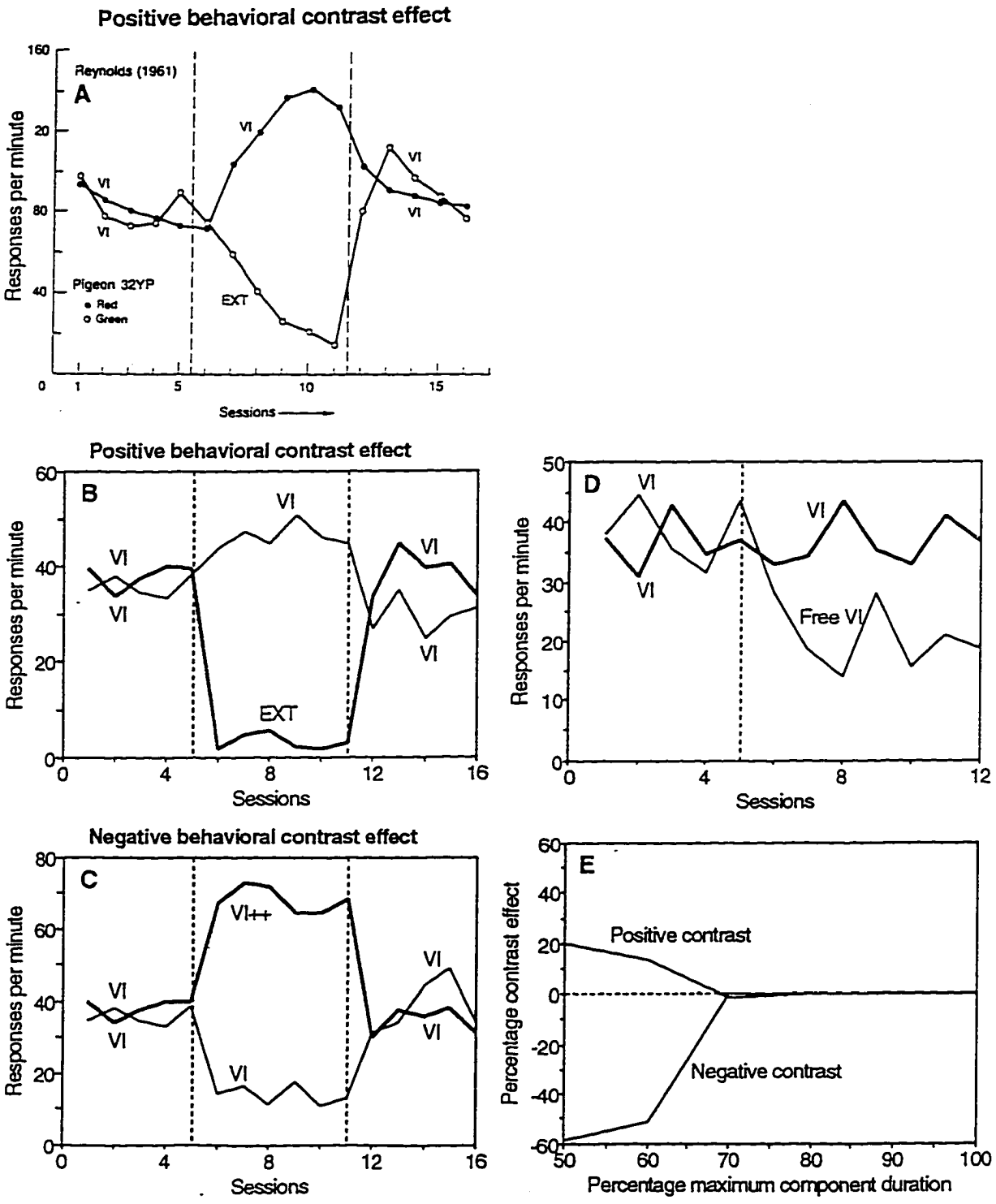
Theoretical analysis of successive contrast effects shows that both positive and negative contrast should be sensitive to *changes* in reinforcement contingency, not just to the simple accumulation of reinforcement. The critical theoretical variables to explain successive contrast are behavioral excitation (a measure of reinforcement underprediction) and behavioral inhibition (a measure of reinforcement overprediction).

Behavioral contrast

Experimental data. Reynolds (1961) showed that in multiple VI-VI schedules in which each variable-interval component is signaled by a different stimulus, the rate of responding is a function of the relative, rather than absolute, rate of reinforcement. In several experiments, the stimulus on the key was red or green and components of the multiple VI-VI schedule alternated every 100 s. When the schedule in one stimulus is changed from VI to extinction while the reinforcement conditions in the other VI schedule are unchanged, pigeons show a substantial and reliable increase in response rate in the unchanged component (positive behavioral contrast effect - Figure 16A). Reynolds also reported a negative contrast effect when reinforcement rate is increased rather than decreased. For instance, when the schedule in one stimulus is changed from VI 100-s, say, to VI 30-s while the other schedule is unchanged the result is a decrease in response rate in the unchanged VI 100 component.

Simulation results. I simulate a multiple VI 100 VI 100 procedure in which different stimuli signal each component of the multiple schedule, exposing the model to 16 sessions (60 reinforcements per session) in which components are alternated.

Figure 16 *Behavioral contrast effect.* (A) Positive behavioral contrast effect. Pigeons are exposed to a multiple VI schedule (each component signaled by red-green stimuli). After the shift (VI to extinction in one schedule) the response rate in the extinction component (green light) drops to almost zero whereas the response rate in the unchanged component (red light) increases [taken from Reynolds (1961)]. (B) Positive behavioral contrast effect. Simulation results obtained using a multiple VI 100-s VI 100-s schedule. (C) Negative behavioral contrast effect. Simulation results obtained using a multiple VI 100-s VI 100-s schedule. After the rate of reinforcement in one component of the multiple schedule is improved to VI 30 (VI++), the rate of response in the unchanged VI 100 component (VI) is degraded. (D) Effect of non-contingent reinforcement. When a multiple VI-VI schedule is changed to a multiple VI-VT schedule, the reduction in response rate obtained in the second component (VT) is not accompanied by an increase in response rate in the first component. (E) Effect of component duration. The degree of schedule interaction is greater the shorter the components of the multiple schedule. Simulations were run to measure the magnitude of positive contrast (VI 100 - VI 100 was changed to VI 100 - extinction) and negative contrast (VI 100 - VI 100 was changed to VI 30 - VI 100) while varying the percentage-maximum component duration (10,000 time units).



For the first and last 5 sessions, both components are the same. For the middle 6 sessions, one component is shifted to extinction. Figure 16B illustrates the simulation results. This figure shows that the response rates in the two components before and after the shift are roughly the same. However, after the shift the response rate in the extinction component (green light) drops to almost zero whereas the response rate in the unchanged component (red light) increases substantially. (It is shown a smaller contrast effect than Reynolds, because the choice of parameter values was a compromise designed to provide a reasonable quantitative fit to a maximum number of operant conditioning phenomena).

The theory explains behavioral contrast as a result of the interplay between behavioral inhibition and response competition. When extinction occurs, the output of the behavioral inhibition unit associated to the green stimulus increases and reduces the corresponding response-strength unit. In these conditions, the alternate response is disinhibited from the mutual interresponse inhibition (response competition) and increases in strength. After extinction ceases the behavioral inhibition unit of the "green" response decays to zero while the VI 100 schedule ensures that the "green" response gradually recovers to the preshift level, inhibiting the "red" response which returns to the previous strength level.

Figure 16C illustrates the simulation results in the negative behavioral contrast effect. This figure shows that after rate of reinforcement in one component (green light) of the multiple schedule is improved to VI 30 (labeled VI++ in Figure 16C) the rate of response in the unchanged VI 100 component (red light) declines. When the VI 30 schedule is shifted back to VI 100 (after 6 sessions) rate of response in both components returns to the preshift level. The theory explains the negative behavioral contrast effect (Figure 16C) as a result of the interplay between behavioral excitation and response competition. After the shift in reinforcement rate occurs, the output of the behavioral excitation unit associated with the green stimulus increases determining the response strength to grow above the preshift level. In these conditions, the alternate "red" response is inhibited at a higher rate than

before the shift, and therefore it decreases in strength. After the VI 30 schedule is replaced with the VI 100 schedule the behavioral excitation unit of the "green" response gradually decreases to the preshift level diminishing the strength of inhibition to the "red" response that gradually returns to the previous strength level (Figure 16C).

Figure 16D illustrates the fact that the reduction in response rate in one component is not the critical variable that determines behavioral contrast. Halliday & Boakes (1971) showed that when a multiple VI-VI schedule is changed to a multiple VI–variable-time (VT: response-independent reinforcers spaced like the VI schedule, or free VI), the reduction in response rate that occurs in the VT component (because food deliveries are no longer response-contingent) is not accompanied by an increase in response rate in the first component: i.e., no positive contrast. To simulate Halliday & Boakes' experiment a multiple VI 100 - VI 100 schedule is applied for five sessions (60 reinforcements per session), followed by a VI 100 - VT 100 schedule for the rest of the sessions. Figure 16D shows that although the rate of responding decreases during the VT schedule, no contrast occurs in the unchanged VI with which the VT is alternated. This happens in the model because free reinforcement delivery somewhat strengthens the o-response (which designate the set of "other" responses) which inhibits the previously contingent response. Furthermore, response competition ensures that the o-response also inhibits the contingent response in first VI 100 component, which prevents positive contrast. The VI-VT sequence is different from the VI-extinction sequence that yields positive contrast. During extinction there is no reinforcement; hence, the o-response cannot increase in strength and inhibit the contingent response. Because of this weak response competition, the contingent response is disinhibited and shows positive contrast in the VI-EXT case but not in the VI-VT case.

One of the most important determinants of behavioral contrast is component duration: the degree of schedule interaction is greater the shorter the components of the multiple schedule (Shimp & Wheatley, 1971; Ettinger & Staddon, 1982:

etc.). As shown in Figure 16E, this effect is true for both positive and negative contrast. Computer simulations were run to measure the magnitude of positive contrast (VI 100 - VI 100 was changed to VI 100 - extinction) and negative contrast (VI 100 - VI 100 was changed to VI 30 - VI 100) while varying the percentage maximum component duration from 50% to 100%. I found that the deviation from the baseline response rates was greatest at short component durations, and qualitatively the effect of component duration is symmetric for positive and negative contrast (although the magnitudes are different). This result is explained by the fact that short component durations determine the increase in response strength in the unchanged component that further increases the strength of response competition. If response competition is stronger the extinguished response is weaker determining a higher level of disinhibition for the response in the VI component that shows positive contrast. Negative contrast is explained in the same way.

The explanation for behavioral contrast supports the view that rate of reinforcement in the adjacent components determines the current response rate (Herrnstein, 1970; Williams, 1983). Consistent with Hinson & Staddon's (1978) views that use behavioral competition as a mechanism for schedule interaction, I found that response competition is critical.

Partial reinforcement extinction effect (PREE)

The PREE has two main aspects: (a) Subjects trained to respond to infrequent reinforcement stabilize at a performance value (e.g., rate of responding) generally lower than subjects trained with more frequent reinforcement. (b) When reinforcement is discontinued (extinction) partially reinforced subjects persist longer in responding than subjects that have been reinforced more frequently, even though they begin extinction responding at a lower rate. The magnitude of the PREE is affected by numerous factors, such as reinforcement probability, reinforcement delay, pattern of reinforcement, reinforcement size, intertrial interval, and length of training. The model correctly predicts how the magnitude of the PREE is affected

by all of these factors, except reinforcement patterning, which requires additional assumptions. In the present study I show the effects of reinforcement probability and length of training.

Effect of reinforcement probability

Experimental data. Probability of reinforcement during acquisition is one of the most important determinants of the PREE (Weinstock, 1958; Bacon, 1962; Kacelnik et al., 1987). For instance, in Kacelnik et al.'s experiment starlings chose between two "foraging patches" in which food was delivered according to either rich or lean probabilistic schedules. There were two comparison groups: in both the lean schedule was 0.08; one group was reinforced with probability 0.25 (rich schedule) and the other one with probability 0.75 (rich schedule). The main conclusions are (Figure 17A): (a) Eventually the rich patch is always preferred close to 100%, but (b) the preference for probability .25 develops more slowly that the preference for probability .75; (c) after the suppression of reinforcement, responses reinforced with probability .25 are more resistant to extinction than responses reinforced with probability .75.

Simulation results. I have simulated a concurrent probabilistic schedule, similar to the foraging situation described by Kacelnik et al. (1987). The probability of the lean side is held at a fixed level (0.08), whereas the probability of the rich side is varied (0.75, 0.33, and 0.25). As shown in Figure 17B, the rate of acquisition is faster when the reinforcement probability on the rich side increases. After 1,000 responses recorded on both sides during acquisition, the reinforcement is extinguished. The proportion of responses on the rich side is calculated for each block of 20 responses throughout the simulation. According to the simulation in Figure 17B, responses reinforced with probability .25 are more resistant to extinction than responses reinforced with probability .33 that are more resistant to extinction than responses reinforced with probability .75, even though during acquisition the proportion of responses on the rich side is directly related to reinforcement prob-

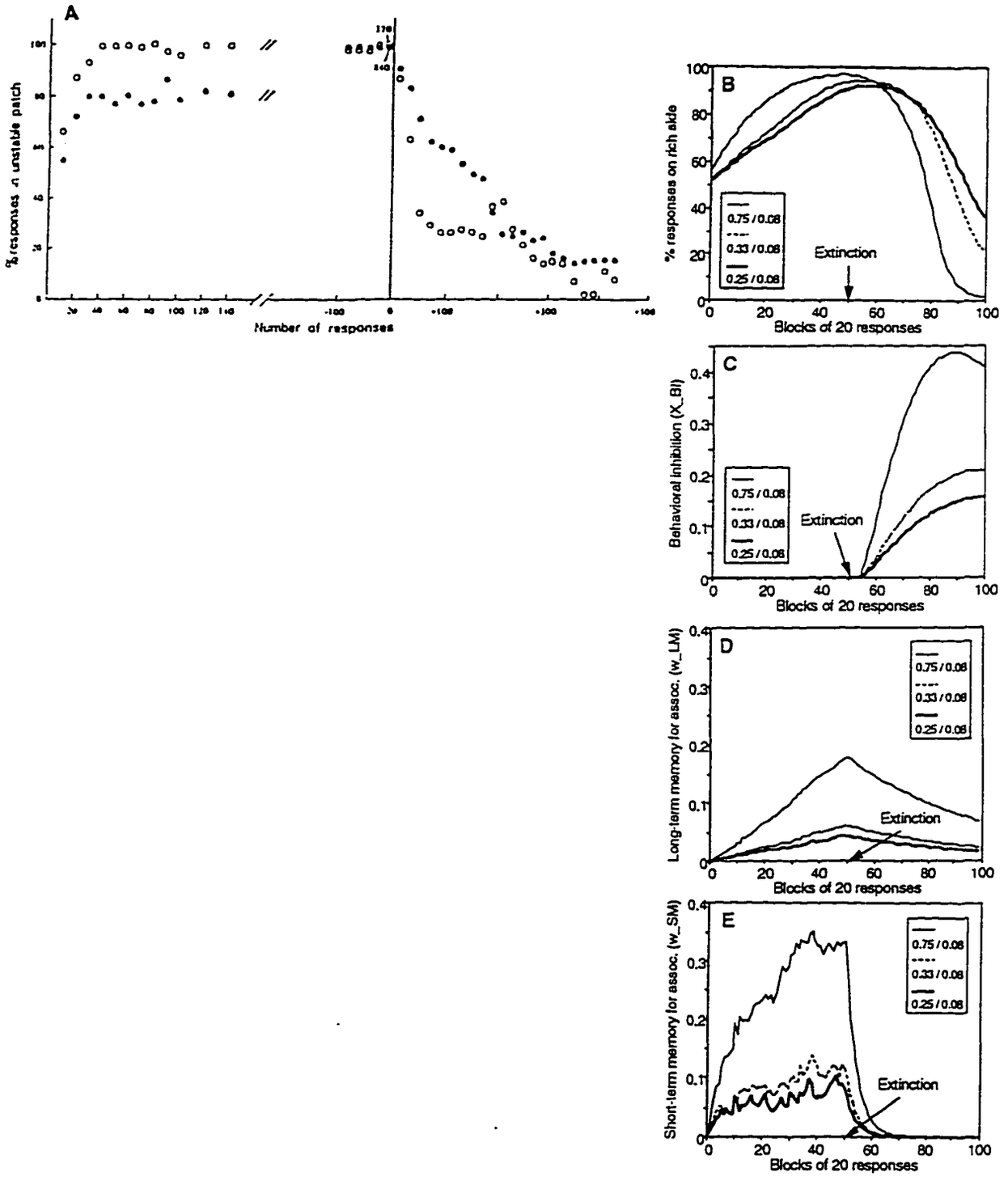
ability.

The fact that the PREE is stronger when the reinforcement is presented less frequently during acquisition can be explained in the following way: the richer the acquisition schedule the higher the level at which both short and long-term memory traces for associations (w_{SM} and w_{LM}) increase. When extinction begins, both w_{SM} and w_{LM} decay, but at different rates: w_{SM} decays faster than w_{LM}, see Figure 17 (panels D-E) and Equations 4, 5. The dynamics of w_{SM} and w_{LM} drive the increase of the behavioral inhibition unit (X_{BI}) via the mismatch between the long and short-term learning expectancy units ($X_{LE} - X_{SE}$). X_{BI} increases during extinction as the difference $X_{LE} - X_{SE}$ becomes greater than 0. Figure 17C shows that the level of the behavioral inhibition unit during extinction is proportional to the reinforcement probability during training. Since the rate at which the response strength is suppressed is controlled by the level of the behavioral inhibition unit, there will be more resistance to extinction with more intermittent schedules (Figure 17B).

In short, the richer the PRF schedule the higher the learning expectancy during acquisition, and thus the higher the level of inhibition (due to X_{BI}) that X_{RS} receives during extinction. Therefore, more intermittency during training determines more resistance to extinction when the frequency of reinforcement alone is varied.

There are very few theories that explain the PREE; one of the most successful is Daly & Daly's (1982) DMOD model, mainly developed to explain classical conditioning effects, incorporating Rescorla & Wagner's (1972) ideas and the major assumptions of frustration theory (Amsel, 1962). DMOD accounts for trial-by-trial changes and asymptotic values in discrete-trial PREE, but cannot account for other aspects of resistance to extinction presented in the following subsections. Unfortunately, most theories of extinction in operant conditioning are verbal, and the very few existing dynamic models have serious difficulties in dealing with the PREE.

Figure 17 *Partial reinforcement extinction effect - influence of reinforcement probability.* (A) Experimental data [taken from Kacelnik et al. (1987)] in which starlings chose between two "foraging patches" in which food was delivered according to either a rich or a lean probabilistic schedule. The lean schedule was 0.08: one group was reinforced with probability 0.25 (rich schedule) and the other one with probability 0.75 (rich schedule). Two aspects are important: (i) Responses reinforced with probability .25 are preferred less rapidly than responses reinforced with probability .75; (ii) During the absence of reinforcement responses reinforced with probability .25 are more resistant to extinction than responses reinforced with probability .75. (B) Simulation results obtained in conditions similar to the original experiment. The acquisition is slower with the decrease in the absolute value of reinforcement probability. Less resistance to extinction with the richer schedule of reinforcement during acquisition. (c) Behavioral inhibition. During extinction, the long-term learning expectancy decreases slower than the short-term learning expectancy causing the increase of behavioral inhibition to a level proportional to the reinforcement expectancy. (D) Long-term memory for associations. During acquisition w_{LM} increases to a level proportional to the actual reinforcement probability. The rate of increase is slower and the curve is smoother compared to w_{SM}. (E) Short-term memory for associations. During acquisition w_{SM} increases to a level proportional to the actual reinforcement probability. After 1,000 responses, the reinforcement is extinguished and w_{SM} decays at a high rate.



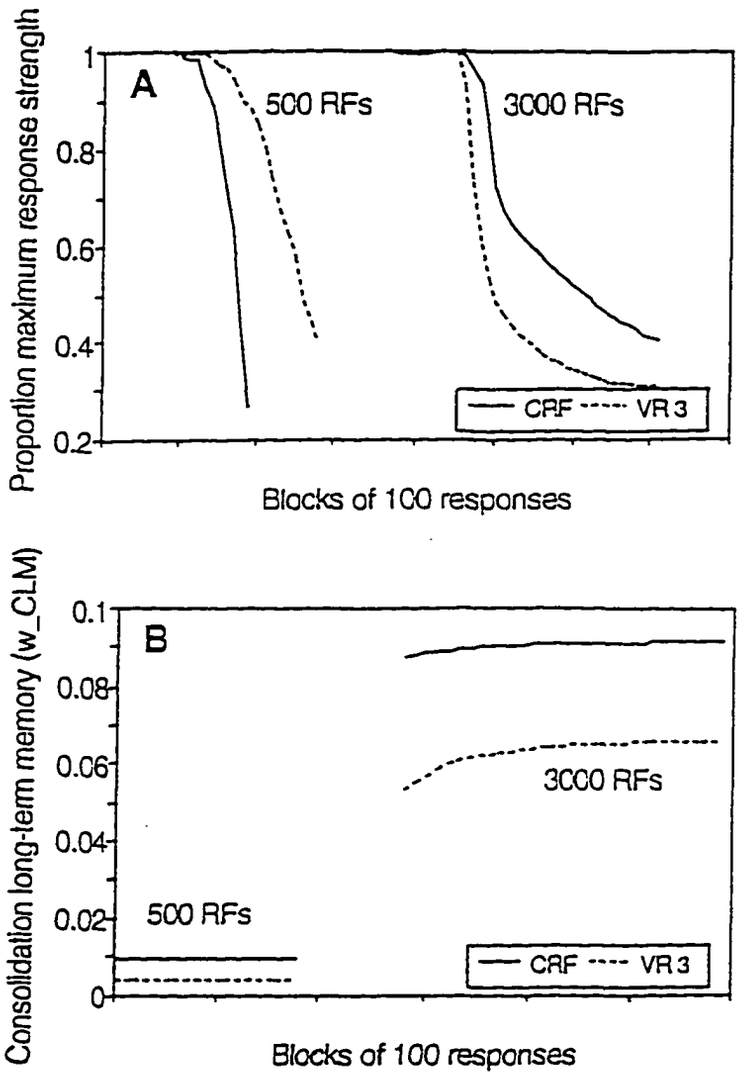
For instance, the cumulative-effects model (Davis et al., 1993) is able to obtain PREE only in special conditions. Staddon (1993), in a discussion of rate-sensitive habituation, has suggested that PREE and successive negative contrast effects both reflect a multiple-time-scale process, but provided no comprehensive account. The analysis shows that the resistance to extinction effect is intrinsic to the model, because of the multiple time scales that underlie response-reinforcement associations. PREE is explained in a way which is not too different from the generalization hypothesis (which is, admittedly, verbal and leaves "generalization" pretty much undefined). However, the problem with the generalization hypothesis, in addition to the fact that it is verbal, is that it does not explain the "reversed PREE" result (discussed in a subsequent section).

Effect of length of training on resistance to extinction

Experimental data. Nevin (1988) has reported that the slope of the extinction curve is flatter after continuous reinforcement than after intermittent reinforcement, suggesting greater rather than less resistance to extinction. This result, obtained in free-operant situations, seems to contradict the classic PREE, i.e., that resistance to extinction is inversely related to rate of reinforcement. But what Nevin actually shows is that it is only when subjects are given extended training that extinction after PRF is faster than after CRF. Apparently the PREE depends on length of training (number of reinforcements), a variable that is often neglected in secondary accounts of the phenomenon.

Simulation results. The theory can accommodate both the PREE and the reversed PREE by utilizing the historical information represented by the consolidation long-term memory variable. If training is not extensive, the effect of consolidation long-term memory is negligible, and the amount of resistance to extinction is controlled by behavioral inhibition that increases proportional to the strength of response-reinforcement associations that are formed during acquisition.

Figure 18 *Partial reinforcement extinction effect - influence of length of training.* (A) Resistance to extinction as a function of length of training (number of reinforcements). The classic PREE (resistance to extinction is a negative function of the rate of reinforcement) is obtained when the reinforcement training is not extensive (e.g., 500 reinforcements). With extended training (e.g., 3,000 reinforcements) the PREE reverses. (B) Consolidation long-term memory is an increasing function of reinforcement probability and number of reinforcements.



As more response-produced reinforcements accumulate during extended training, w_{CLM} increases, gradually masking the inhibitory effect produced by X_{BI} during extinction. The level at which w_{CLM} increases is strongly dependent on the effective reinforcement probability, so that the consolidation long-term memory facilitates the response more strongly as the rate of reward approaches CRF, thus producing more resistance to extinction in the CRF than the PRF group. Figure 18 shows simulation of an acquisition-extinction situation in which the length of training was varied between two groups, one receiving a CRF schedule, and the other receiving a VR 3 schedule. Two distinct situations are analyzed: extinction after 500 reinforcements and extinction after 3000 reinforcements. Figure 18A shows that as the number of reinforcements increases, the difference in resistance to extinction between continuously and intermittently reinforced responding also increases, replicating Nevin's (1988) reverse-PREE finding. Figure 18B shows the cause of this effect: the increase in consolidation long-term memory follows the increase in the length of training.

I am not aware of any other model which can accommodate both the PREE and the reversed PREE. The fact that the theory incorporates history-sensitive variables which encode the response-reinforcement association strength throughout training is essential. The hypothesis is that the operant response is the result of two competing forces: one is suppressive (behavioral inhibition) and tends to reduce the response strength during nonreinforcement, and the other one is facilitatory (consolidation long-term memory) and tends to counterbalance the suppressive effects by exciting the response strength unit, at a much slower time scale, as a result of previous reinforcement.

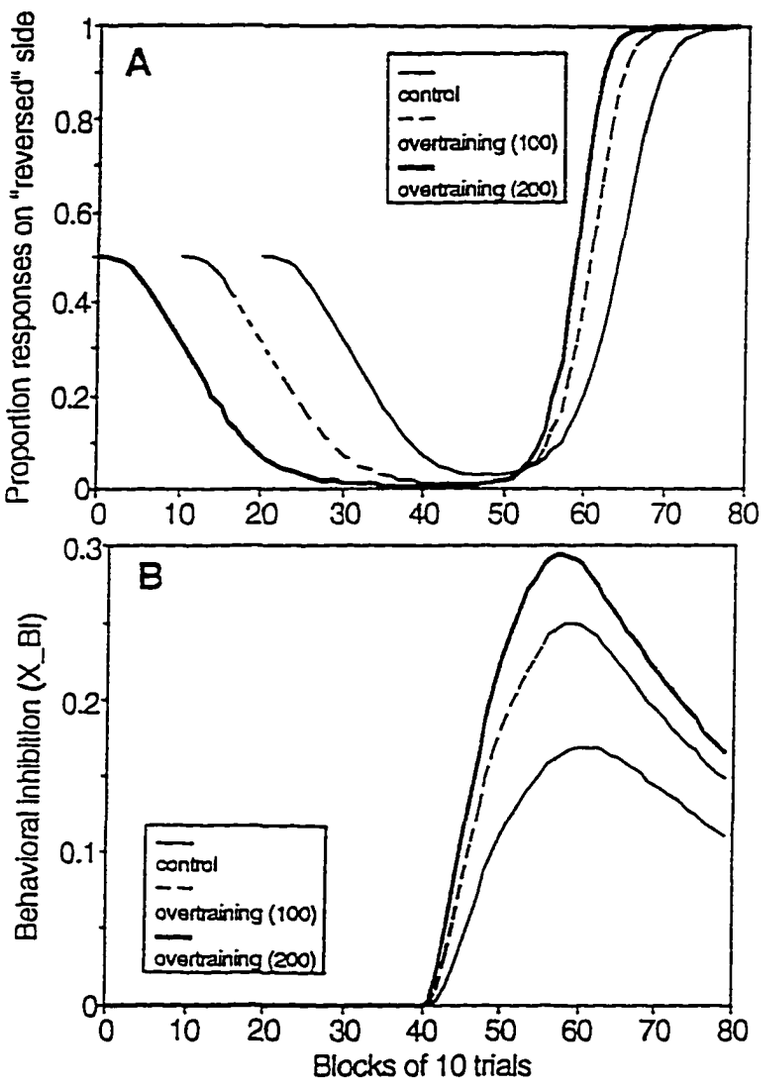
Overtraining Reversal Effect (ORE)

Experimental data. The ORE is the fact that overtrained subjects often learn the reverse discrimination faster than control subjects. For instance, if one group is reversed immediately, one after 100 trials of overtraining, and one after 200 trials

of overtraining, the speed of reversal is proportional to the amount of overtraining (Reid, 1953; Hooper, 1967). Mackintosh (1969) reported a significant ORE in animals given a large reward, and a much smaller effect in animals given a small reward. Reid (1953) and Capaldi & Stevenson (1957) showed that reversal was facilitated by a moderate number of trials beyond a certain criterion, and was markedly facilitated by more extended overtraining. However, these latest results are inconsistent with Sperling (1970), who used a criterion of learning different from Reid and Capaldi & Stevenson, a fact which made Mackintosh (1974) "forcibly impressed by the inconsistency of the effect and its reluctance to submit to any simple analysis" (p. 603). Although the necessary and sufficient conditions for the ORE are still obscure, there are some regularities: (a) ORE is stronger if large reward is used; (b) ORE magnitude increases with the amount of overtraining.

Simulation results. ORE is hard to explain by theories that assume that more training is equivalent to more strength of stimulus-response association. I explain ORE qualitatively, mainly because of the lack of detailed performance information on the effect (see Mackintosh, 1974, for a review). Computer simulations are presented in a discrete-trial choice situation in which the reward is given after each correct response on one of two alternatives. There are three groups: the control group is reversed after 200 trials, and the other two groups are reversed after 300 and 400 trials, respectively. Figure 19A shows that speed of reversal is highest in the group reversed after 200 overtraining trials, and it is minimum in the control group. The theory attributes the faster reversal to the right, say, with extended training to the left to the fact that extended training strengthens short and long-term memories for response-reinforcement associations which then determine an increased level of learning expectancy. Therefore, when the left response is in extinction, since behavioral inhibition increases at a rate controlled by the long-term learning expectancy, the left response in the overtrained group is suppressed at a higher rate compared to the control group (which received less reinforcers during acquisition).

Figure 19 *Overtraining reversal effect.* (A) Simulations are presented in a discrete-trial choice situation in which the reward is given after each correct response on one of two alternatives. There are three groups: the control group is reversed after 400 trials, and the other two groups are reversed after 500 and 600 trials, respectively. The speed of reversal is highest in the group reversed after 200 overtraining trials, and it is minimum in the control group. (B) Behavioral inhibition increases proportionally to the amount of overtraining.



Thus, the right response is released faster from interresponse inhibition (faster reversal to the right). Figure 19B shows that behavioral inhibition increases at a rate controlled by the amount of overtraining.

The theory also predicts that the ORE will be stronger with larger reinforcements because bigger reward means higher learning expectancies which means a stronger behavioral inhibition signal during extinction, and therefore faster reversals. In fact, no effect is obtained if the reward magnitude was too small. in agreement with a number of findings (e.g., Hooper, 1967; Mackintosh, 1969). The theory also predicts that the magnitude of the ORE is not monotonically related to the number of overtraining trials. That is, because of the nonlinearities in the model, I expect that control animals trained to a certain criterion reverse only slightly more slowly than animals given an additional of N trials before reversal, but the control animals should reverse much more slowly than animals given $2N$ trials before reversals. These predictions are actually consistent with the findings of Reid (1953) and Capaldi & Stevenson (1957). I am not aware of any quantitative model which is able to generate ORE, although the cumulative-effects model (Davis et al., 1993) is able to produce it in some special conditions.

Spontaneous recovery

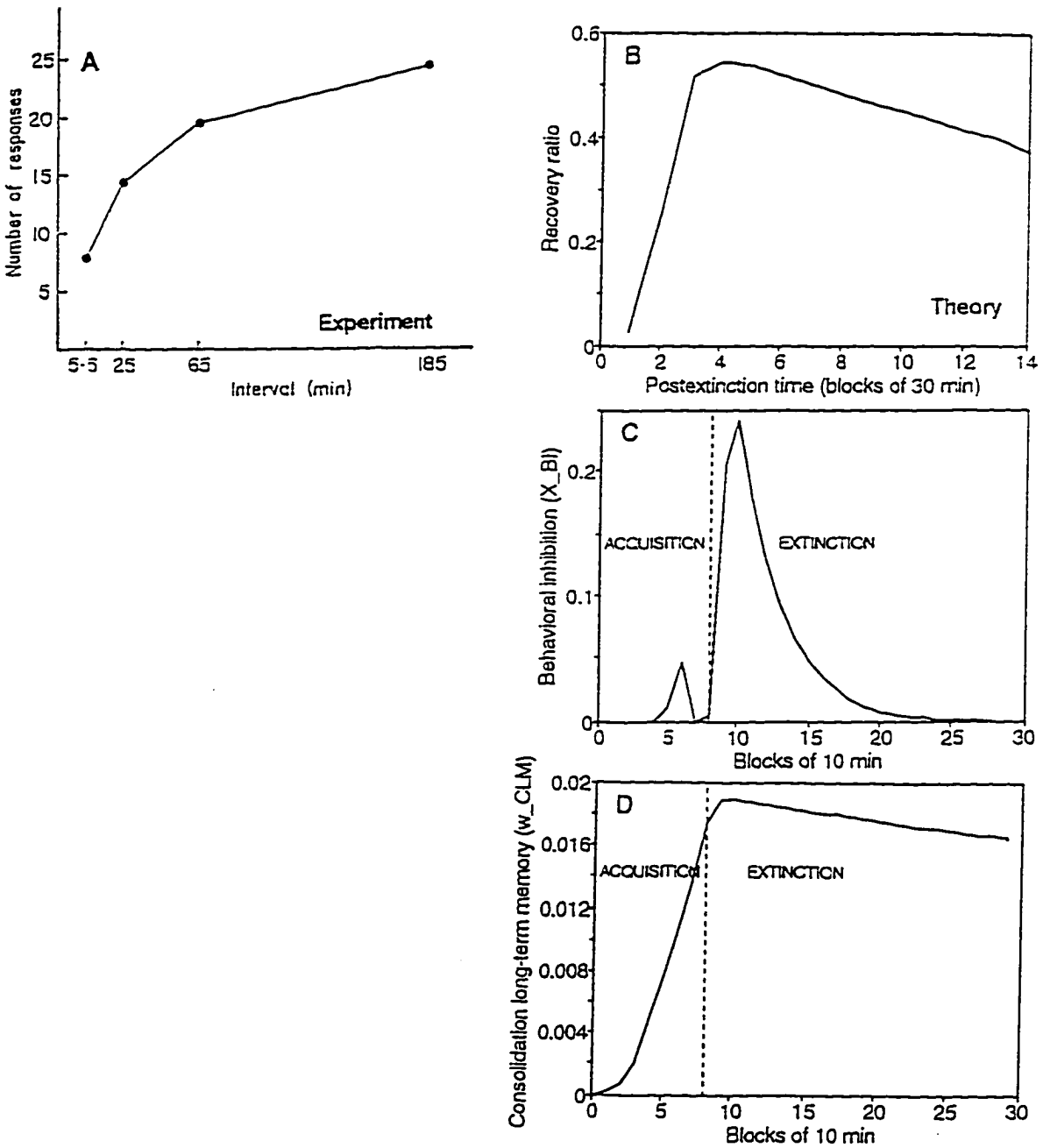
Changes in performance may occur over an interval of time when the subject is not exposed to reinforcement contingency, and even when the subject is not exposed to the experimental situation at all (inter-session time). One interesting instance of such "spontaneous" change is spontaneous recovery (SR) after extinction when after some interval since the termination of the prior extinction the subject's initial pattern of response increases as a function of the inter-session time. I will deal with discrete-trial single-response recovery studies first and free-operant choice later.

Spontaneous recovery in discrete-trial single-response situations

Experimental data. The quantitative estimate of spontaneous recovery is the mean recovery ratio, defined as proportion of responding relative to the total amount of responding during acquisition. Spontaneous recovery can be described by two distinct phases: (a) within hours following extinction recovery reaches a maximum estimated at roughly 40% of the initially acquired response pattern. (b) after the recovery maximum is reached, the amount of recovery dissipates slowly in time such that responding can be observed for many days following the end of extinction (Mackintosh, 1974; Robbins, 1990). Figure 20A illustrates only the first stage of SR in which the "spontaneous" responses reach a maximum after roughly 3 hours since the offset of extinction.

Simulation results. To test the operation of the theory for SR I have simulated an acquisition session of 80 min where responses are paid off with probability 0.1 (VR 10 schedule), followed then by extinction. Figure 20B shows the amount of recovery as a function of the postextinction interval. Each point on the curve represents the proportion of responses in blocks of 30 min postextinction time units. Consistent with experimental data, two phases can be distinguished: after the response becomes fully extinguished it recovers within an interval (estimated roughly as 120 min) to about 55% of the initial level of responding, followed by a relatively slower decay that can last for many hours. The explanation for the dynamics of recovery is the following: in the absence of concurrent alternatives, the response is influenced by two types of activation (cf. Equation 10): one facilitatory, $(w_1 + w_{CLM})X_{BE}(1 - X_{RS})$, expressing the influence of long-term learning expectancy (I do consider the effect of consolidation long-term memory), and one suppressive, $-\alpha_3 X_{BI} X_{RS}$, expressing the inhibitory influence exerted as a consequence of nonreinforcement. The superposition of these two effects generates the profile depicted in Figure 20A.

Figure 20 *Spontaneous recovery in discrete-trial single-response situations.* (A) Number of responses as a function of postextinction interval [adapted from Mackintosh (1974)]. (B) Recovery ratio (ratio between the postextinction response strength and the baseline) as a function of postextinction time. Two phases can be distinguished: after the response becomes fully extinguished it recovers within an interval (estimated roughly as 150 min) to about 50% of the initial level of responding, followed by a relatively slower decay that can last for many hours. (c) Behavioral inhibition decays with the time since extinction, gradually releasing the response strength from inhibition. (D) Consolidation long-term memory (w_{CLM}) slowly increases during acquisition. After extinction begins w_{CLM} decays slowly.



During extinction all the memory traces for associations decrease. The decay process happens at different rates, depending on each memory trace, i.e., w_{SM} decays fast, w_{LM} decays slower, and w_{CLM} decays even slower. It was previously shown that due to the decay of the memory units the behavioral inhibition unit increases to a maximum during extinction (until the mismatch between the long and short-term learning expectancy becomes 0) and then decays spontaneously. Figure 20C shows that behavioral inhibition decays with the time since extinction and consequently inhibits the response strength unit at lower levels. At the same time, the behavioral excitation unit facilitates the response strength unit. The level of facilitation is controlled by the connection strength $w_1 + w_{CLM}$, where w_1 is fixed and w_{CLM} is variable.

Immediately after extinction, the activity of the behavioral inhibition unit is more effective than the facilitatory effect induced by behavioral excitation (activity modulated by the slowly varying consolidation long-term memory). As the postextinction time goes on and behavioral inhibition decays faster than the consolidation long-term memory (Figures 20C-D), the response strength is slowly released from inhibition and, due to the facilitation from behavioral excitation, it increases to a maximum. After behavioral inhibition becomes 0, the response strength decreases slowly (the effect of recovery can last for days), as it relies only on the slowly decaying consolidation long-term memory, thus explaining the second portion (after the recovery ratio reaches a maximum) of the curve depicted in Figure 20B.

The role of discriminative stimulus is to trigger the operant behavior after a prolonged absence of reinforced training (during spontaneous recovery). In these conditions, after all the memory traces decay to zero, the consolidation long-term memory is the only variable which encodes information about the history of reinforcement. When the animal is reintroduced into the experimental box the discriminative stimulus switches from 0 to 1 and triggers a burst of responding (in the case of a nonzero w_{CLM}). The responses thus generated read-out the consolidation long-term memory and increase the response strength at a rate controlled

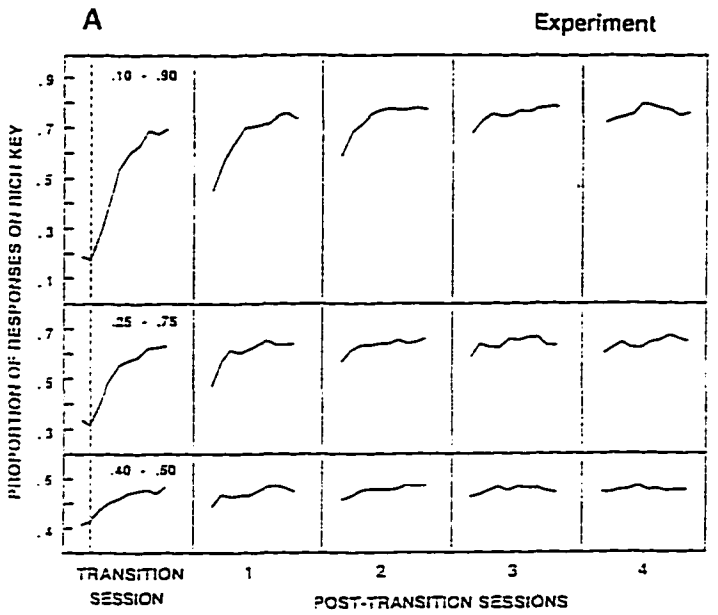
by w_{CLM} (the higher w_{CLM} the higher the recovery rate). Because these "recovery" responses are generated during extinction (when $X_{LE} > X_{SE}$), the behavioral inhibition unit is able to become active and to suppress the pattern of responding.

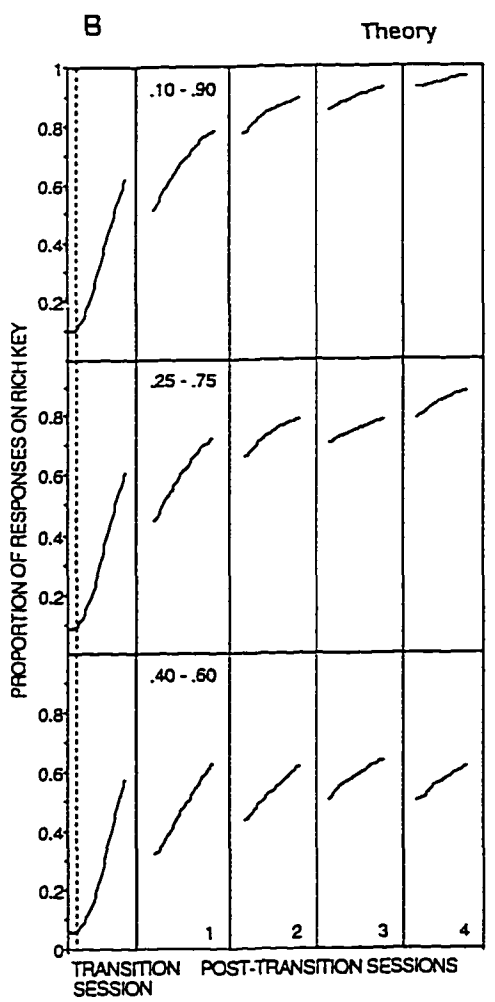
Spontaneous Recovery in Choice

Experimental data. By comparison with discrete-trial procedures, there has been little research on spontaneous recovery in free-operant studies. Recently, Mazur (1995) has performed a between-session spontaneous recovery study in which he observed that response proportions at the beginning of a new session were not the same as those at the end of the previous session. There was a tendency to revert toward response proportions of earlier sessions. Specifically, in a choice experiment in which reinforcement is assigned by a single VI 30-s, during a pretransition session the probability of assigning a reinforcement to the left key was .10, .25, .40, .60, .75, or .90. The pretransition session was followed by one transition session in which the same probability was in effect for the first 6 min, and then the probabilities for the two keys switched for the rest of the transition session and for four posttransition sessions. Mazur reports that at the start of each posttransition session the proportion of responses to the rich key was lower than at the end of the preceding session, this spontaneous recovery being largest in the first posttransition session and then becoming progressively smaller in subsequent sessions (Figure 21A).

Simulation results. Figure 21B illustrates the simulation results. In the interval between the end of the transition session and the beginning of the first posttransition session both short and long-term expectancies decay to their rest level, such that when the posttransition session begins the response to the lean key relies only on the slowly decaying consolidation long-term memory (which encodes the past experience, before the switch in the reinforcement probabilities). Therefore, there is a recovery of the "lean" response toward the level before the switch.

Figure 21 *Spontaneous recovery in choice.* (A) Reinforcement is assigned by a single VI 30-s, during a pretransition session the probability of assigning a reinforcement to the left key was .10, .25, and .40. The pretransition session is followed by one transition session in which the same probability was in effect for the first 6 min, and then the probabilities for the two keys switched for the rest of the transition session and for four posttransition sessions. At the start of each posttransition session the proportion of responses to the rich key was lower than at the end of the preceding session, this spontaneous recovery being largest in the first posttransition session and then becoming progressively smaller in subsequent sessions [adapted from Mazur (1995)]. (B) Simulation results obtained in conditions similar to the original experiment.





The recovery response for the "lean" key, stronger than the response for the "lean" key at the end of the transition session, gives a response proportion (for the rich key) which is lower than at the end of the transition session. However, since the consolidation long-term memory decays in the intervals between the posttransition sessions, the "recovery" response is weaker with each posttransition session and thus the SR magnitude diminishes.

Mazur explains spontaneous recovery in his choice experiment by a simple "global" hypothesis, namely that responding at the start of each new session is a function of responding during several previous sessions, not just during the previous session. This view describes the global cumulation process that underlies the cumulative effects model (although that model cannot explain Mazur's data because it has no provision for stimuli or time), and is also consistent with the consolidation long-term memory that encodes remote reinforcement experience in the present model. The present model, in addition, predicts that the magnitude of spontaneous recovery will be lowest in the first posttransition session, rather than largest, as Mazur found, if the intervals between posttransition sessions are made shorter (e.g., at the end of the transition session use short "probe" sessions separated by 30-min intersession intervals) - because the course of SR is determined by time rather than trials in the model. This experiment does not appear to have been done.

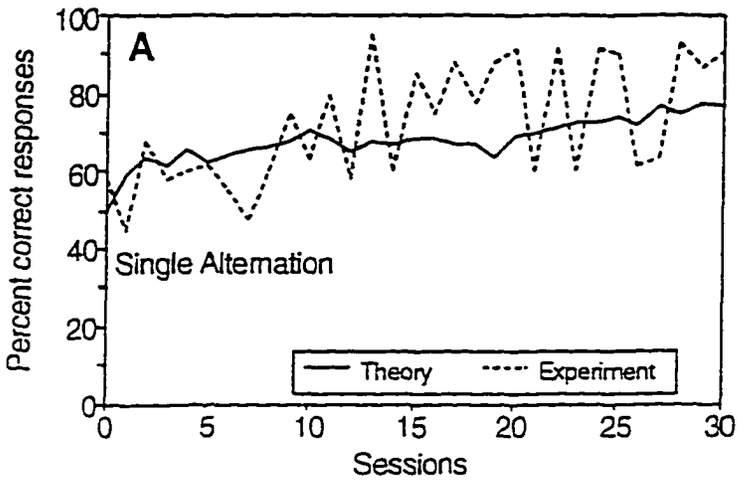
There are a few theories of spontaneous recovery, most verbal (Pavlov, 1927; Capaldi, 1967, 1971; Rescorla & Wagner, 1972; Mackintosh, 1974), and only very few quantitative (e.g., Estes, 1955; unfortunately, Estes' model is at a molar level, and therefore is unable to explain the real-time mechanism for recovery). The cumulative-effects model (Davis et al., 1993) can only account for a particular form of spontaneous recovery, i.e., regression, that is encountered in choice experiments when, in extinction, there is a reversion to an earlier preference, despite the fact that this alternative is no longer rewarded.

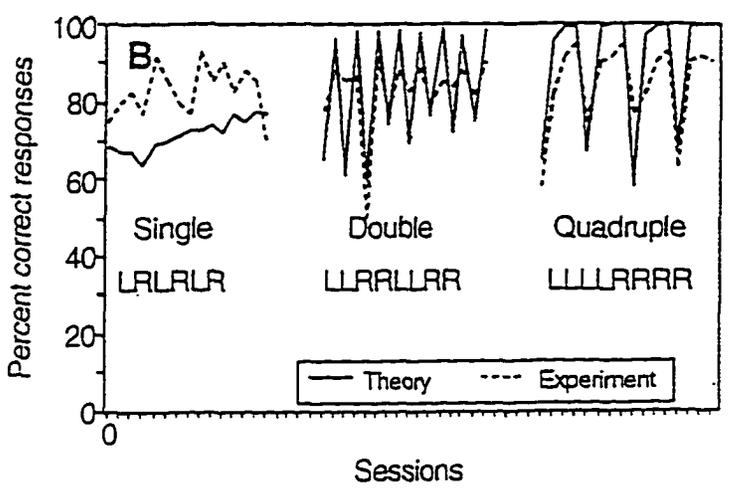
Serial Reversal Learning

Experimental data. Reversal learning involves alternating extinct-reinforcement contingencies between two choices. After a certain number of trials, subjects show progressive improvement in performance. In Davis & Staddon's (1990) experiments pigeons were exposed to successive daily discrimination reversals and reversals in blocks of two and four days in the two-armed-bandit situation. Subjects had to choose between two response keys, left (L) and right (R), each reinforced according to probabilistic schedules (1/8). The "hot" key varied from day to day in a regular fashion (LRLR, LLRR, or LLLLRRRR). Figure 22A shows that pigeons improve their performance across days of reversals. A result that could not have been easily predicted beforehand is that pigeons reverse faster after exposure to daily reversals than to reversals in blocks of two days, and they reverse faster after exposure to reversals in blocks of two days than to reversals in blocks of four days (Figure 22B) - a result similar to the data on frequency of preference-switching of Mark and Gallistel, discussed earlier.

Simulation results. I simulate the successive daily reversals problem (LRLR case) and reversals in blocks of two (LLRR case) and four (LLLLRRRR case) days. In all three cases there were 30 training sessions, with 60 reinforcements per session. In the case of daily reversals, during each day responses to a different key are reinforced with probability 1/8. In the case of reversals in blocks of two/four days, during each series of two/four consecutive days, responses to a different key are reinforced with the same probability, 1/8. Figure 22A shows that the percent correct responses for one of the alternatives, e.g., left, improves over sessions of training. Improvement in performance is due to the slowly varying consolidation long-term memory, w_{CLM}, which increases at a faster rate than it decays (see Appendix).

Figure 22 *Serial reversal learning.* (A) Improvement in discrimination performance across successive discrimination reversals. The reinforced key (each response is paid off with probability 1/8) changes every session (LRLR - single alternation); solid line: simulation results; dashed line: experimental results. (B) Discrimination-reversal performance in successive reversals (single - LRLR) and reversals in blocks of 2 (double - LLRR) and 4 (quadruple - LLLLRRRR) sessions; solid line: simulation results; dashed line: experimental results. Reversals are faster after exposure to single alternations than to double alternations, and faster after exposure to double alternations than to quadruple alternations [experimental data adapted from Davis & Staddon (1991)].





According to Equation 10, as training progresses w_{CLM} causes X_{RS} to increase faster each day and to inhibit the alternative response at a higher level. At the same time, learning expectancy increases over sessions contributing to the increase in behavioral inhibition that suppresses the incorrect responses (responses in extinction). Both effects contribute to increases in performance, i.e., faster reversals.

The theory also predicts (Figure 22B) faster performance when reversals are performed on a daily basis compared to reversals in blocks of two days, and faster performance when reversals are performed in blocks of two days compared to reversals in blocks of four days. The explanation relies on the dynamics of consolidation long-term memory. For instance, in the case of double alternations the reinforcement provided for the alternative that experiences two consecutive days of acquisition (LLRR) is then extinguished for the next two days, and therefore w_{CLM} will decay more in extinction for the LLRR pattern than for the LRLR pattern. Thus, when the previously extinguished alternative is reinforced again, w_{CLM} starts increasing from a lower level compared to the LRLR case because of the two consecutive extinction sessions during which w_{CLM} slowly decreased. In a similar way it is explained why reversals are faster in double alternations compared to quadruple alternations (Figure 22B).

The explanation for the reversal learning effects presented above is not too different from the explanation given by the authors of the original experiment (Davis & Staddon, 1990). Their mechanism relies on the properties of the ratio between the number of reinforcements for responding at one side and the total number of responses on that side (Davis et al., 1993), with reinforcements and responses calculated over the whole training history. Therefore, the critical variable is the fact that the state of the model is sensitive to the whole training history. In the same vein, in the theory the key variable for explaining reversal effects is the consolidation long-term memory, a memory trace that encodes the effects of reinforcement over a large time window.

Effects of context on stimulus preference

A problem which is still under debate is how the animal uses reinforcement history to solve the context-dependent choice between two alternatives which were previously experienced in conjunction with other alternatives. In a series of recent experiments, Williams & Royalty (1989) and Belke (1992) question the validity of context-free theories of choice that rely on reinforcement probability as the sole variable controlling choice behavior.

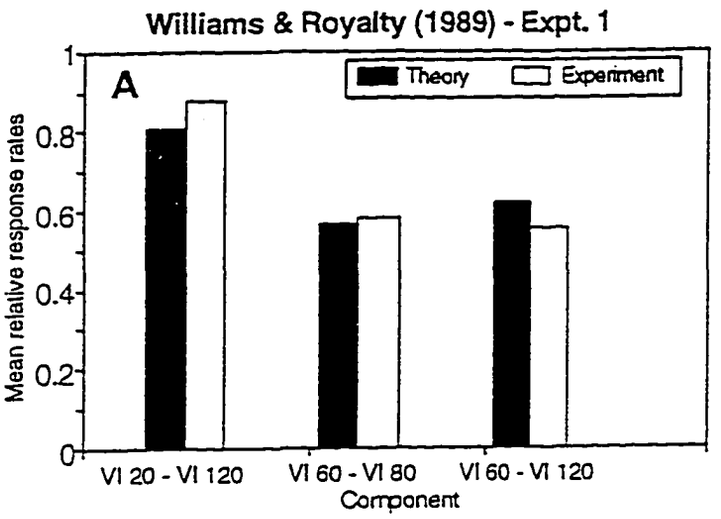
Experimental data. In an attempt to test the melioration hypothesis (Herrnstein & Vaughan, 1980), Williams & Royalty (1989) showed that contrary to melioration absolute, not local, rates of reinforcement determine preference. Their experimental procedure (Experiment 1) comprised two alternating multiple-schedule discrete-trial components. During the first component, pigeons had to choose between a VI 20-s and a VI 120-s schedule; the second component was a choice between a VI 60-s and a VI 80-s schedule. Each component was signaled by a different stimulus. As expected, pigeons approximately matched response and reinforcement ratios during components 1 and 2 of the multiple schedule. However, when the stimulus that signaled the VI 60-s schedule was pitted against the stimulus that signaled the VI 120-s schedule, with extinction in effect for both alternatives, Williams & Royalty found that the VI 60-s stimulus (paired with a lean schedule during training, hence associated with a low obtained reinforcement rate) was slightly preferred over the VI 120-s stimulus (paired with a rich schedule during training, hence associated with a high obtained reinforcement rate). This result contradicts the melioration hypothesis according to which the stimulus paired with the schedule that delivers the highest obtained reinforcement rates during acquisition (VI 120-s schedule) should have been preferred.

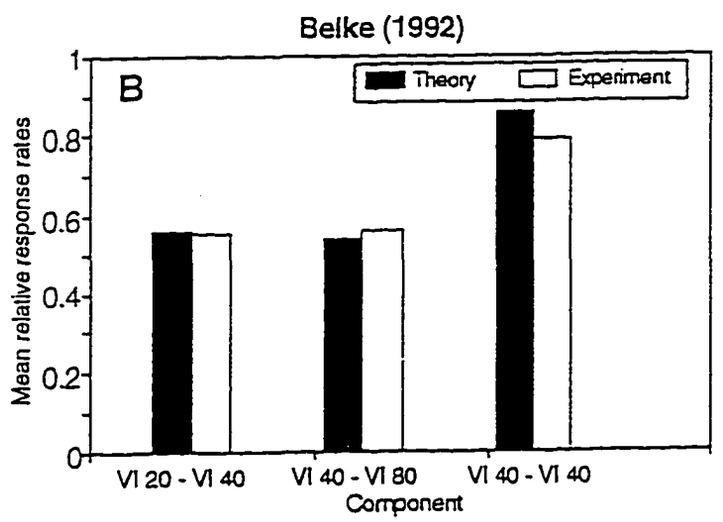
Simulation results. Model prediction, along with the experimental data, are shown in Figure 23A. During acquisition the model mean relative response rates are 0.81 (experimental 0.87) in the first component (VI 20 - VI 120) and 0.56 (experimental 0.575) in the second component (VI 60 - VI 80), approximately equal

to the matching ratios. In agreement with Williams & Royalty's data, the theory correctly predicts that the VI 60-s schedule is preferred over the VI 120-s schedule by a factor of 0.62 (experimental 0.554, measured by averaging the results of 5 of the subjects). This result is explained in the following way. The two stimuli that signal the two components of the multiple schedule enter both response units corresponding to the left and the right keys, both stimuli being contingent on the highest reinforcement probability. Therefore, their contribution to the behavioral excitation units is the same in the case of both alternatives. However, this is not the case with the left and the right responses. Since the richer schedule is always associated with one key, say left, the learning expectancy of the left response increase faster than the learning expectancy of the right response. Therefore, when the stimuli signaling the VI 60-s schedule and the VI 120-s schedule are pitted together the left responses show a higher level of behavioral excitation compared to the right responses. In this way, the response strength unit for the left key (reinforced according to the VI 60-s schedule) increases to a higher level and inhibits the response strength unit for the right key (reinforced according to the VI 120-s schedule), and therefore the VI 60-s schedule is preferred.

Experimental data. Belke (1992) showed that neither absolute nor local rates of reinforcement account for context-dependent preference. He showed that a difference in the absolute rates of reinforcement between concurrent VI-VI schedules is not necessary to produce a distinct pattern of preference. Preference reflects the previous history of reinforcement associated with the choice stimuli. The experimental procedure comprised two alternating multiple-schedule with concurrent VI components. Pigeons were trained on a multiple concurrent VI 20-s VI 40-s, VI 40-s VI 80-s, each schedule being associated with distinct response keys and signaled by a different color stimulus. I label these four schedules A, B, C, D. During training pigeons approximately matched response and reinforcement ratios during both components of the multiple schedule.

Figure 23 *Effects of context on stimulus preference.* (A) Mean relative response rates in each component of the Williams & Royalty (1989) Experiment 1 (two alternating multiple-schedules): first component - VI 20-s VI 120-s schedule; second component - VI 60-s VI 80-s schedule; test component - VI 60-s VI 120-s schedule. Filled bars: simulation results; empty bars: experimental data. (B) Mean relative response rates in each component of Belke's (1992) experiment (two alternating multiple-schedules): first component - VI 20-s VI 40-s schedule; second component: VI 40-s VI 80-s schedule; test component: VI 40-s VI 40-s. Each schedule is associated with a distinct response keys and is signaled by a different color stimulus. Filled bars: simulation results; empty bars: experimental data.





However, when the stimuli that signal the two VI 40-s schedules (B-C) were pitted against each other, with extinction in effect for both, Belke found that the stimulus associated with the VI 40-s schedule previously paired with the VI 80-s schedule (C) is strongly preferred over the stimulus associated with the VI 40-s schedule previously paired with the VI 20-s schedule (B), despite the fact that the rate of obtained reinforcement and the relative response ratios for the two VI 40-s schedules were similar during training. Belke interpreted this result, that is pigeons prefer the stimulus associated with the schedule that was preferred in previous pairings over a stimulus associated with the same schedule but had not been preferred in previous pairings, in terms of the context-dependent differential properties of stimulus preference.

Simulation results. Model prediction and experimental data are shown in Figure 23B. During training the model mean relative response rates on the VI 40-s schedule in each component approximately matched the experimental response rates on the VI 40-s schedule (roughly 0.56 in each component). In agreement with Belke's data, the theory correctly predicts that the VI 40-s schedule that was previously paired with the VI 80-s schedule (c) is preferred over the VI 40-s schedule that was previously paired with the VI 20-s schedule (B) by a factor of 0.85 (experimental 0.79). This result is explained as a global effect of response competition that determines the preference pattern during training. I label responses reinforced by schedules A, B, C, and D as R_A, R_B, R_C, and R_D. Since during training schedules B and C deliver the same rate of reinforcement, the behavioral excitation units that correspond to R_B and R_C will increase to the same value facilitating the corresponding response strength units. However, since the rate of reinforcement provided by schedule A is four times higher than that provided by schedule D, R_B is under strong local inhibition from response R_A, whereas R_C is under weak local inhibition from response R_D. The differential inhibition strength reflects into the level of the two response strength units, i.e., $X_{RS}^B < X_{RS}^C$. Therefore, when the stimuli associated with schedules B and C are pitted against one

another the local inhibitory effects transfer at a global level and, given the same rate of reinforcement for both schedules B and C, response competition ensures a higher level for X_{RS}^{C} than for X_{RS}^{B}, and therefore more frequent responses for the stimulus associated with schedule C (previously associated with VI 80-s) than responses for the stimulus associated with schedule B (previously associated with VI 20-s).

The effects of context on stimulus preference presented in this section can also be explained by models which rely on the dynamics of time allocation between competing alternatives. For instance, Gibbon et al. (1988) assume that subjects acquire "knowledge" of the actual interreinforcement times associated with each response alternative, calculated with respect to those intervals in which a given alternative is available, a principle used to explain Williams & Royalty's (1989) results. In a subsequent analysis, Gibbon (1995) explains Belke's (1992) results by assuming that subjects allocate time to each alternative in a two-phase process. The first phase is a decision process for deciding between two alternatives by sampling from exponential memories and choosing the more reinforcing of the two samples. The second phase is a memory sampling process performed at a rate controlled by the overall arousal level (reinforcement rate) in the training context. This two-phase process makes subjects in Belke's experiment slower in deciding in the probe signal from the leaner pair (VI 40-s VI 80-s) than in the probe signal from the richer pair (VI 20-s VI 40-s), resulting in the surprising time-allocation preference for the VI 40-s schedule, previously paired with the VI 80-s schedule, pitted against the VI 40-s schedule previously paired with the VI 20-s schedule. Gibbon's theory shares with the present model the idea that the time-allocation for each alternative is determined by the overall reinforcement rate. Thus, whereas Gibbon hypothesizes that the memory sampling rate is proportional to the overall rate of reinforcement, I use the fact that response competition is stronger at high overall rates of reinforcement to account for the fact that the subject allocates less time to the alternative paired with a rich schedule. This explanation (arousal

level increases the level of response competition) is also consistent with Gibbon's (1995) new finding, i.e., the puzzling time-allocation preference for a VI 40-s schedule, previously paired with a VI 80-s schedule, pitted against a VI 20-s schedule previously paired with a VI 40-s schedule.

In a critique of incremental "response strength" models, Williams (1994) suggests that we should rule out models which rely solely on reinforcement probability as the fundamental controlling variable, and embrace "representational" views according to which "the subject acquires veridical knowledge of the frequency of reinforcement associated with the different response alternatives, and then on the basis of some choice rule allocates its behavior accordingly (p. 707)." Williams also suggests that we need to "embellish probability-based models with additional assumptions about the nature of the response unit". The theory offers a way to reconcile stimulus strength models with "representational" theories by giving response units an internal structure (e.g., response-reinforcement associations, learning expectancy, response competition, etc.) that provides a substrate for context, one of the basic assumptions being that the performance at the level of each response depends on the response configuration at the level of all response alternatives.

Experimental analysis

In this section I use the model to test the idea that learning is driven by *both* local and global processes. The theory presented in the previous sections has been applied to a wide range of operant conditioning phenomena at different time scales. The major interest in this section is to investigate whether the theory can suggest new experiments to help understanding the nature of choice behavior. Under stable environmental conditions it is well established that in probabilistic reinforcement schedules responding eventually fixates on the richest choice alternative that provides the maximum reinforcement (e.g., larger magnitude, higher probability of occurrence, shorter latency). However, what alternative will be preferred when both reinforcement probabilities are made equal? Many studies indicate

that preference is distributed equally between the two choices, whereas at least one other study shows that choice can fixate on one randomly determined alternative (Horner & Staddon, 1987). I claim that the pattern of preference with two equal alternatives cannot be predicted without knowing the reinforcement history. Unfortunately, I have not been able to find an appropriate study that addresses this issue.

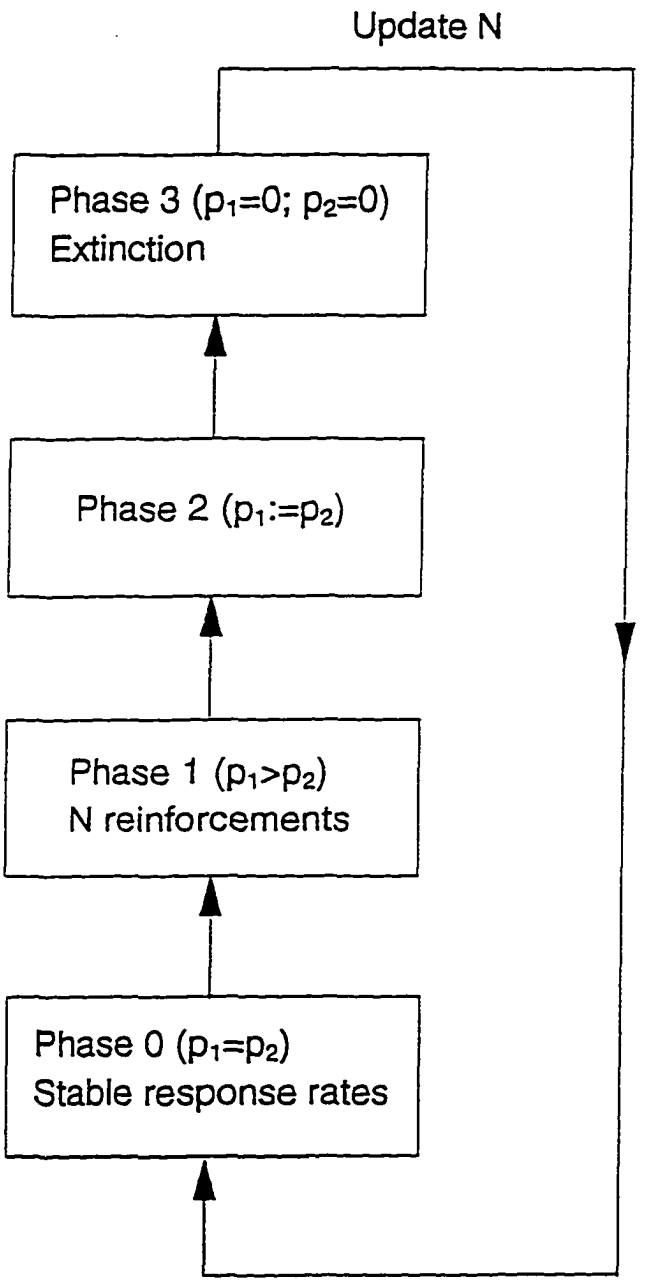
A critical test of the theory is to use model predictions to understand how historical information (amount of training) biases the pattern of preference when the environment is destabilized. Specifically, how the historical information is being used to choose between two equal alternatives that did provide reinforcement at different rates in the past.

Method

The following experiment compares the performance of 4 experienced hungry pigeons (85% free-feeding weight) in a situation in which historical information (amount of training) is used to evaluate the pattern of preference in a two-armed bandit paradigm (responses on two alternatives are reinforced with different probabilities, or concurrent VR-VR schedule). The subjects were trained in a two-key pigeon chamber in which each key had a different color. A dim houselight provided general illumination, except during reinforcement, when it was replaced by a light over the hopper. The procedure was implemented and data collected by a 386 computer located in an adjacent room.

The experiment consists of a four-phase cycle, three acquisition phases and one extinction phase (see Figure 24 for a schematic diagram). The four phases are: Phase 0 – concurrent VR-VR interdependent schedule in which both reinforcement probabilities are equal to the mean of the reinforcement probabilities in the next phase. A reinforcement is offered for pecking the left/right key only if the previous reinforcement has been offered for responding to the opposite side.

Figure 24 *Schematic diagram of the experimental procedure.* p_1 and p_2 represent reinforcement probabilities for responding at two keys. Phase 0 – interdependent reinforcement schedule in which $p_1 = p_2$; Phase 1 – concurrent VR-VR schedule in which $p_1 > p_2$ for N reinforcements; Phase 2 – concurrent VR-VR schedule in which the previous phase is maintained for 10 reinforcements after which both probabilities are made equal to the lowest probability ($p_1 := p_2$); Phase 3 – Extinction ($p_1 = p_2 = 0$). The cycle continues with a new condition being tested (e.g., number of reinforcements, N).



This phase is necessary to obtain stable unbiased response patterns (each acquisition session consists of 60 reinforcements). Phase 1 – two-key concurrent VR-VR situation with probabilities p_1 and p_2, where $p_1 > p_2$. This phase consists of a variable number N of sessions (1 or 5, with 60 reinforcements per session). Phase 2 – one transition session which begins with the same reinforcement conditions as Phase 1. After 10 reinforcements the probability of the rich side becomes equal to the reinforcement probability on the lean side, i.e., $p_1 := p_2$. The transition session is followed by one identical session ($p_1 = p_2$), but without the 10-reinforcement pretransitory component. Phase 3 - extinction sessions that last until an interresponse time greater than 10 minutes. At the end of Phase 3 a new condition is tested (i.e., different number of reinforcements in Phase 2 and/or a new pair of reinforcement probabilities, p_1 and p_2) by exposing the same animal to Phases 0, 1, 2, and 3 presented in this order.

Results

I measured the percentage responses on the "lean" key (responses on this key were always reinforced with probability p_1) in the last session of Phase 1 and in the transition session of Phase 2. The number of reinforcements in Phase 1 was varied from 60 (1 session) to 300 (5 sessions) across conditions. Figure 25 shows the real-time pattern of preference in all subjects during the last session of Phase 1 ($p_1 = 0.15$, $p_2 = 0.05$) and during the transition session of Phase 2 ($p_1 = p_2 = 0.05$) when the number of reinforcements in Phase 1 is 60. In Phase 1 the percentage of responses on the "lean" key indicates that the rich side is greatly preferred. However, in Phase 2, after both reinforcement probabilities are made equal, the response shows strong preference (over 60% in all subjects over the last 200 responses collected on both keys) for the previously lean side.

Figure 25 *Experimental results (N=60 RFs).* Data are from four pigeons exposed to a concurrent VR - VR schedule in which the "lean" and the "rich" keys are reinforced with probabilities $p_1 = 0.15$ and $p_2 = 0.05$ (Phase 1). After 60 reinforcements in Phase 1, the reinforcement probability for the "rich" key becomes equal to the reinforcement probability for the "lean" key (Phase 2). Data is presented in parallel for all four subjects (0027, 0869, 0883, 1346). Average response percentages over the last 10 blocks in Phase 2 are calculated. Responding in Phase 2 shows that following the shift in reinforcement probability preference gradually changes toward the "lean" key (preference over 60% in all four subjects). The vertical dashed line denotes the approximate block when Phase 2 was applied. The figure shows the percentage of responses in the last session of Phase 1 and in the transition session of Phase 2.

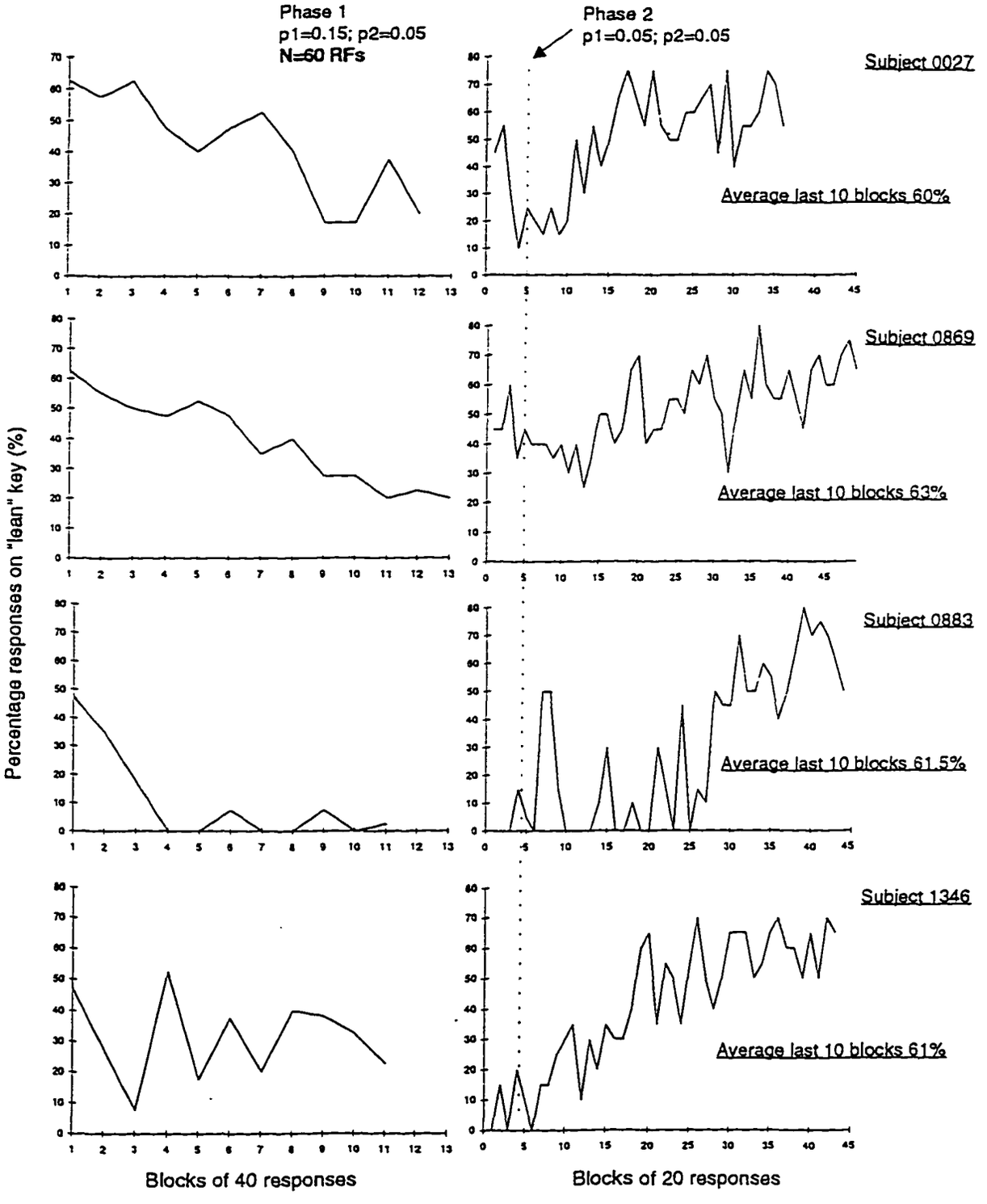
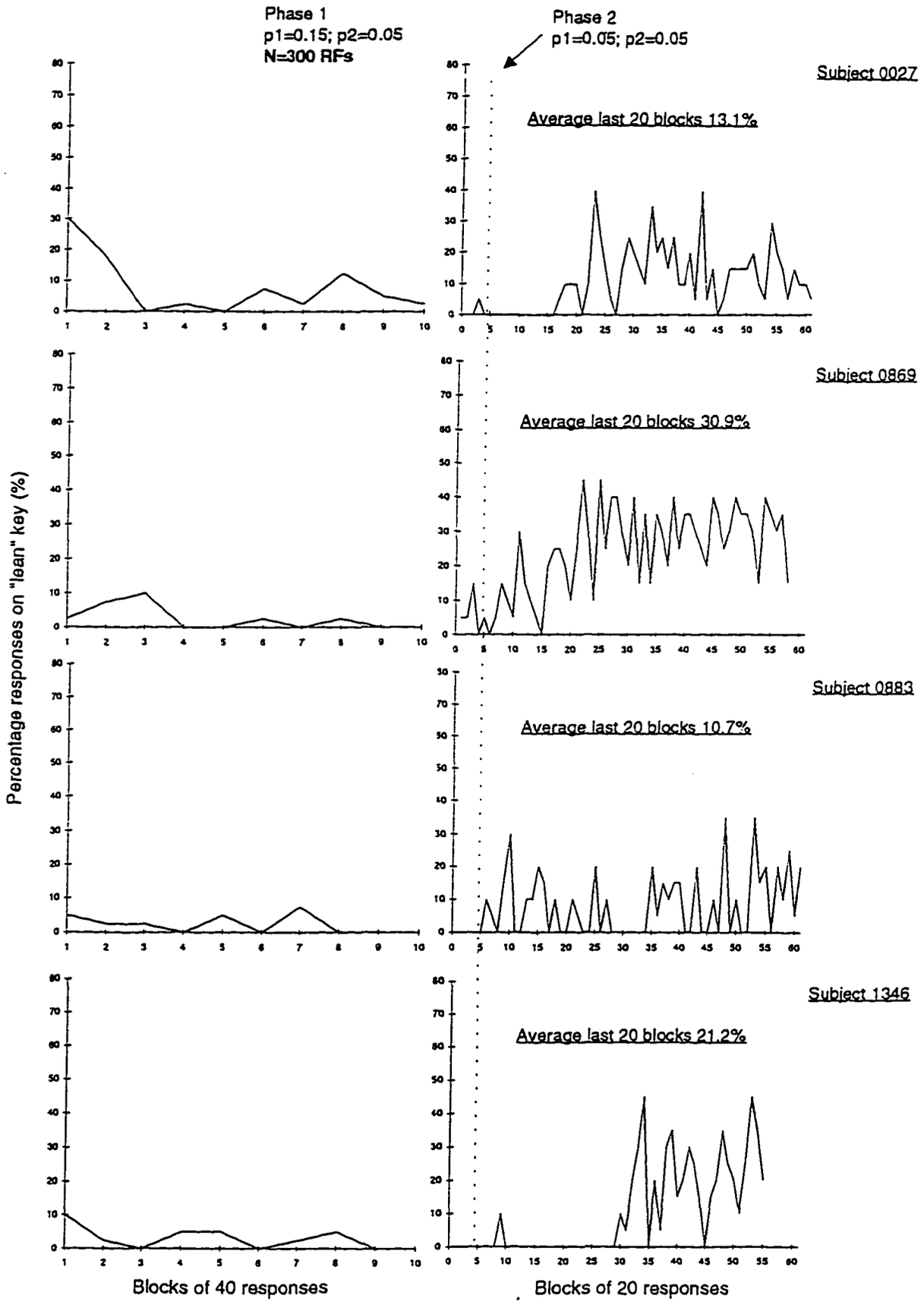


Figure 26 *Experimental results (N=300 RFs).* Data are from four pigeons exposed to a concurrent VR - VR schedule in which the "lean" and the "rich" keys are reinforced with probabilities $p_1 = 0.15$ and $p_2 = 0.05$ (Phase 1). After 300 reinforcements in Phase 1, the reinforcement probability for the "rich" key becomes equal to the reinforcement probability for the "lean" key (Phase 2). Data is presented in parallel for all four subjects (0027, 0869, 0883, 1346). Average response percentages over the last 10 blocks in Phase 2 are calculated. Responding in Phase 2 shows that following the shift in reinforcement probability the "rich" key remains preferred (preference over 65% in all four subjects). The vertical dashed line denotes the approximate block when Phase 2 was applied. The figure shows the percentage of responses in the last session of Phase 1 and in the transition session of Phase 2.

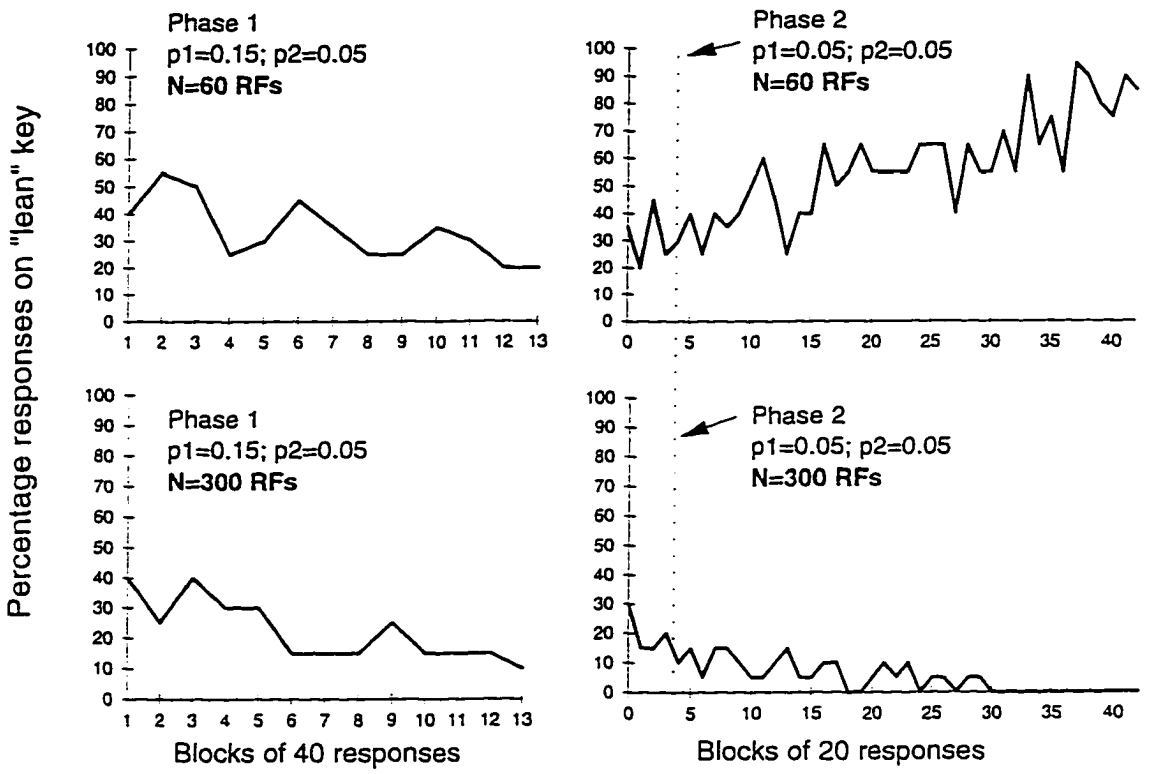


If the number of sessions in Phase 1 is extended from 1 to 5 sessions (300 RFs) the number of responses to the lean key decreases as the preference becomes more exclusive (Figure 26 – Phase 1). In these conditions, when pigeons switch to Phase 2 they continue to prefer the previously rich side (Figure 26 – Phase 2), despite the fact that both reinforcement probabilities are made equal, thus obtaining a completely opposite result from the situation in which only 60 reinforcements are offered in Phase 1.

These surprising results are reliable and they have been obtained in all of the subjects. During Phase 2 data was collected in blocks of 20 responses (rather than 40 as during Phase 1) in order to detect any variation in the distribution of responses between the two keys. The pattern of preference illustrated in Figure 25 (Phase 2) is a transient effect. It typically disappears in the second session of Phase 2 when all pigeons distributed their responses equally between the two keys (50% preference – data not shown). Also, data from Phase 0 and Phase 3 is not shown because of the predictibility of these results (equal response proportions in Phase 0 and extinction in Phase 3).

The simulation results (Figure 27) indicate good quantitative fits to experimental data. Phase 0 has been simulated by starting with equal response proportions to the two keys and identical parameter values associated with the two response units. Phase 3 has been simulated by exposing the model to prolonged extinction until no response was recorded. The results illustrated in Figure 27 are explained in the following way. After 60 reinforcements in Phase 1, the decline in reinforcement probability for the rich side causes a mismatch between the short and long-term expectancies for the "rich" alternative. Since short-term expectancy decays faster than long-term expectancy, $X_{LM} - X_{SM} > 0$ and therefore the response on the "rich" side is suppressed. In these conditions (reinforcement overprediction) responding on the "rich" side drops below the level that would be normally expected when the reinforcement probability is 0.05, and the competition between responses ensures that preference shifts toward the "lean" side.

Figure 27 *Simulation results.* Percentage of responses on the "lean" key was calculated in the following conditions: Phase 1 ($p_1 = 0.15; p_2 = 0.05; N = 60RFs$); Phase 2 ($p_1 = 0.05; p_2 = 0.05; N = 60RFs$); Phase 1 ($p_1 = 0.15; p_2 = 0.05; N = 300RFs$); Phase 2 ($p_1 = 0.05; p_2 = 0.05; N = 300RFs$). Each phase is described in the text. All the parameters used here have the same values as in the previous simulations (see *The dynamics of operant conditioning*), except for $\alpha_{12} = 10^{-4}$ and $\alpha_{13} = 10^{-2}$.



However, if Phase 2 occurs after a longer exposure to Phase 1 (5 sessions), the contrast effect shown in Figure 27 (60-RFs condition) disappears and preference remains stable for the previously rich side. This result can be explained as an influence of the facilitatory effect of the consolidation long-term memory (w_{CLM}) that is able to cancel the suppressive influence of reinforcement overprediction. Relatively extensive reinforcement training in Phase 1 ensures a high level of the consolidation long-term memory for the "rich" key, an effect which tends to mask the suppressive influence of the mismatch between long- and short-term expectancies, and animals react to the change in reinforcement probability as during the one-session Phase 1.

The effects illustrated in Figures 25 and 26 cannot be accommodated by a simple stimulus strength theory. For instance, if one assumes that the effect of reinforcement is cumulative, then both before and after the occurrence of the shift in reinforcement probability preference will always stabilize to the "rich" side. The reason is a higher response strength value for the rich side, even after the shift. In contrast to these results, the theory predicts that upon the occurrence of the change in reinforcement probability the preference should shift from the side which previously provided reinforcement with probability 0.15 to the side which previously provided reinforcement with probability 0.05, despite the fact that both short and long-term obtained reinforcement rates for this side are at a lower level than those of the other side. Furthermore, the present model correctly predicts that the effect shown in Figure 25 is reversed after extended training (Figure 26), a result which could be hard to accommodate by other theories of operant conditioning.

The theory-guided results presented here have serious implications for understanding the allocation of preference in operant studies. The main finding is that short- and long-term memory processes coexist and their interaction is crucial. Local models would predict preference for the "lean" side during the transition session (some local models have been shown capable to produce rebound effects as

illustrated in Figure 25, Phase 2), independent on the previous training experience. whereas global models would put more weight on the long-term reinforcement history and therefore they would always predict preference for the "rich" side. Only a model that considers both short and long-term processes, such as the multiple time-scale model presented here, would be able to explain the data presented here, as well as other hard-to-explain phenomena in the operant literature (see for example Mark & Gallistel, 1994 and 1995: Nevin, 1988).

Discussion

I hypothesize that the process of conditioning involves the formation of associations between competing operant responses (and environmental stimuli) and the reinforcement. These associations are used to build stimulus- and response-specific short and long-term learning expectancies. The operant response is controlled by the ensemble behavioral excitation and behavioral inhibition which integrates the mismatch between long and short-term learning expectancy. The theory has provisions for historical effects of reinforcement: with extended training the efficacy of response control by the expected reinforcement increases via slow changes in the consolidation long-term memory variable.

Five principles are crucial to the model: (1) Learning dynamics reflect the operation of processes at at least two different time scales (short and long). (2) A key feature of conditioning is formation of estimates of the degree of association between responses and reinforcement and stimuli and reinforcement. (3) Aggregate response-reinforcement and stimulus-reinforcement associations are defined as *learning expectancy*. (4) A *consolidation long-term memory* incorporates sensitivity to the whole reinforcement history. (5) The operant response is controlled by comparison between expected and experienced events. This idea can be contrasted with the ideas advanced by other theories of conditioning which assume that the goal of learning is to accurately predict the actual or the future reinforcement levels, e.g., Klopf (1988), Sutton & Barto (1990), Schmajuk & DiCarlo (1992).

All associations in the model are built as leaky-integrator units, a concept widely supported in the modeling community in both psychology (e.g., Grossberg, 1982: Davis et al., 1993; Schmajuk & DiCarlo, 1992) and neuroscience (e.g., Stemmler et al., 1995; Douglas et al., 1995). Gallistel (1990) has raised a number of objections to integrator models: that a single rate parameter cannot cover all time scales and that "they conflate four unrelated aspects of the problem: what the animal learns, how rapidly it learns it, how rapidly it forgets it, and its uncertainty regarding the correctness of what it has learned" (p. 377). He goes on to conclude that "a representational account is called for" (p. 382). The present model, and other multi-unit models (e.g., Staddon & Higa, 1996), show that these are not criticisms of integrator models in general, but only of the single-unit model. Simple units, appropriately arranged, can do as well as any "representational" scheme, with the added virtue of providing a testable dynamic process.

The present theory uses associations which change continuously in strength in reaction to external events such that at the end of training a new stable state is reached. The stability of new states goes along with partial destabilizing of the previous stable states (e.g., competing responses decrease in strength). During extinction association strengths gradually decrease and the corresponding response loses its stable properties (i.e., rate of increase and asymptotic level). But this "forgetting" is not total. For instance, the slowly varying consolidation long-term memory allows it to encode events over extended time intervals, despite the total "discharge" of the short and long-term memory for associations. The learning mechanism allows each association strength to be moved by the learning process from any state to any other state via reversible state transitions. Although conditioning increases all association strengths, whereas extinction decreases them, these changes are reversible and are fully controlled by the reinforcement conditions. Because of the different decay time constants for the short and long-term associations, behavioral inhibition (a measure of reinforcement overprediction) acts to suppress the response via fixed inhibitory connections. This process is one way to resolve

the long-standing controversy about whether extinction decreases the strength of excitatory associations or strengthens competing inhibitory associations. In the model, extinction only decreases the strength of excitatory associations, which then has an effect on response inhibition.

As it was mentioned earlier, a few other dynamic theories have been proposed to account for the dynamics of conditioning. I conclude by saying a bit more about how the present theory relates to some of the alternatives.

A key aspect of the theory presented here is that the operant response is controlled by interplay between expected (long-term) and experienced (short-term) reinforcement. This principle has been applied for the first time to operant conditioning, but a similar idea has been used previously by some models for classical conditioning. For instance, Daly & Daly's (1982) model (DMOD), Klopf's (1988) drive-reinforcement model (D-R), Sutton & Barto's (1990) temporal differences model (TD) and Schmajuk & DiCarlo's (1992) model (S-D), share with the present theory the idea that learning is driven by the mismatch between expected reinforcement and current events. But besides the completely different implementation of the "mismatch" idea, the encoded event is quite different. The present theory encodes the mismatch between short and long-term learning-expectancy units that represent an aggregate measure of the associations between responses and discriminative stimuli and the reinforcement. Other theories use other kinds of mismatch: (a) between the actual value of the US, either appetitive or aversive, and the sum of three types of stimulus-specific associative strengths: approach, avoidance, and counterconditioning (DMOD), (b) the absolute difference between the US and the aggregate prediction of all stimuli present, including the US (S-D model), (c) changes in stimulus level associated with changes in response level computed at consecutive time steps (D-R model), or (d) successive predictions of reinforcement (TD model). All these theories (a-d) have in common the assumption that the basic conditioning mechanism is built around the mismatch between the US level and the US prediction, or around the mismatch between successive

predictions of the US, whereas the present theory utilizes the mismatch between the *association* between responses or discriminative stimuli and the reinforcement. An important distinction: it is not the reinforcement per se that is basic to operant conditioning, but its association with the response and with contextual stimuli.

Another key feature of the present theory is sensitivity to the history of reinforcement. It is shown how the whole history of reinforcement can be encoded by consolidation long-term memory, a principle useful for explaining phenomena such as spontaneous recovery or improvement in performance across daily reversals. Whole-history sensitivity is not a new concept. It is embodied most directly in the cumulative-effects model (CE: Davis et al., 1993), in the form of separate long-term memo-ries for responses and reinforcements, accumulated during the whole history of training. The CE model assumes that the only link between responses and the reinforcement is the ratio between their long-term memories; but the present theory assumes that responses and reinforcement become associated early in processing; consolidation long-term memory reflects the strength of their connection. But nonlocal (and non-associative and non-temporal) principles (like the CE model) are not sufficient to explain all experimental data. The CE model has difficulty with effects of extended training on positive and negative contrast, because the model implies that responses and reinforcements become less effective after long training, whereas the data show that successive negative contrast is stronger after longer pre-shift training. Of course, local principles (like the Bush-Mosteller-type models or Killeen's, 1994, model) are not sufficient either. Local models cannot explain serial-reversal improvement or spontaneous recovery, for example.

One important distinction between the present theory and other dynamic theories of operant conditioning is its comprehensiveness. I have shown that several apparently unrelated phenomena reflect common underlying processes. I start with a few key principles and show that many operant conditioning phenomena can be explained as emergent properties.

Principles such as response competition and multiple time scales (MTS) in event processing have been used previously in different forms to explain limited data sets, but the strength of the present model lies in its capacity to assemble these principles into a coherent mechanism. However, it would be wrong to assume that *any* model which incorporates competition or multiple time scales can explain the data presented here: earlier MTS models, for example, have attempted to deal only with highly restricted data sets, such as rate-sensitive habituation and some aspects of contrast and the PREE. They were not intended to cover the full range of operant conditioning phenomena. The present model is an attempt to go beyond models that deal only with arbitrarily restricted datasets to cover the full range of phenomena in operant conditioning.

Another point is that history-dependence, a property often ignored in operant conditioning studies, and its quantitative measure (number of reinforcements) should be given more careful consideration. The amount of training should be considered as another independent variable, along with reinforcement probability or reinforcement delay. Its reconsideration is potentially beneficial for the field of operant conditioning, and learning in general. It is shown here that a study that examines the role of length of training in operant studies is able to reconcile apparently conflicting experimental results (which in fact are conflicting only because they reflect large variations in the training history).

Finally, the present theory has implications for animal-learning methodology. Most previous dynamic theories for operant conditioning could not deal with long-term historical effects. This led the field of animal learning to use the so-called "between-group" methodology, which requires large numbers of supposedly "naive" subjects, to be used for a limited number of daily sessions and then "sacrificed". The alternative "within-subject" approach, which uses a small number of animals for long-term experiments, declined in popularity because it appeared to be unsuitable for the study of learning, which depends strongly on the animal's training history. A good real-time theory describes exactly how the state of an animal

depends upon its history. Such a theory can be tested using repeated experiments on the same organism. Thus, instead of trying to equate histories by using large groups of naive animals, a few animals with known histories can be used in repeated experiments.

Short and Long-Range Cortical Dynamics: A Model of Orientation-Dependent Context Effects in Primary Visual Cortex

Abstract

A fundamental property of neurons in primary visual cortex, in addition to their selective response to particular oriented bars within their receptive field, is the influence from outside the "classical" receptive field, usually termed surround or context effects. It is generally accepted that context effects depend on the relative contrast between center and surround (Toth et al., 1996; Weliky et al., 1995; Knierim & Van Essen, 1992). However, increasing evidence argues that context effects are also orientation-dependent. It has been demonstrated that the same center stimulus can both strongly suppress and facilitate responses as a function of surround orientation. Unfortunately, although several studies (e.g., Stemmler et al., 1995; Somers et al., 1995) have investigated the involvement of long-range horizontal connections, that link cells with similar orientation preference over large regions of visual space, as the most popular candidate for explaining contrast-dependent context effects, the exact nature of orientation dependency remains unresolved. Here I describe how a large-scale model that incorporates orientation-dependent effects of horizontal connections and cortical feedback can use the interplay between the effect of short and long-range horizontal connections to disinhibit orientation detectors along a direction away from the surround orientation, an effect amplified by the excitatory cortical feedback. The model's properties are utilized to explain several orientation-dependent context effects in primary visual cortex, as well as context-dependent psychophysical effects such as geometrical illusions of orientation (e.g., Zollner illusion).

Introduction

A prominent feature of the connectivity in primary visual cortex (V1) is the existence of long-range horizontal connections (Gilbert & Wiesel, 1979; Rockland & Lund, 1982; Livingston & Hubel, 1984; Martin & Whitteridge, 1984) that have axon collaterals extending for several millimeters within the superficial layers of V1. Since these connections link cells across distinct regions of the visual field they have the potential to underlie influences on cortical neurons from outside their classical receptive field (the classical receptive field is defined as the region in which a given stimulus triggers a suprathreshold response from the cell).

There are several studies that have investigated the nature of the extraclassical receptive field influences (context or surround effects). The main finding is that context effects in V1 are both contrast and orientation-dependent. Thus, the same surround can either facilitate or suppress the response to a preferred stimulus within the classical receptive field depending on the intensity of the center stimulus. Although the surround stimulus alone, or in the presence of a low-contrast center stimulus, is only able to evoke a weak excitatory response, when the center stimulation is made stronger the response is suppressed. These effects of stimulus contrast are highly reliable (e.g., Weliky et al., 1995; Toth et al., 1996) and several theoretical proposals have tried to account for them (Stemmler et al., 1995; Somers et al., 1995.

The orientation dependency of context effects means that the presence of an iso-oriented surround stimulus beyond the classical receptive field suppresses the response to an optimal stimulus within the classical receptive field (Gilbert & Wiesel, 1990; Knierim & Van Essen, 1992; Grinvald et al., 1994). In contrast, stimulating the extraclassical receptive field with an orthogonal or oblique grating in conjunction with optimal stimulation within the classical receptive field has been shown to "supra-optimally" facilitate responses (Sillito et al., 1995; Lennie et al., unpublished data; Gilbert & Wiesel, 1990). In addition, Gilbert & Wiesel (1990) showed that the presence of surround stimuli oriented counterclockwise from the

cell's optimal orientation is able to shift the cell's orientation tuning curve in a direction away from the surround orientation. The shift is reversible, and is fully controlled by the situations in which the surround is either present or absent.

In all of the context-dependent studies that will be discussed here the distance between the center and surround stimulus varies between 1° and several degrees of visual angle. Hence, the most likely candidate to explain the orientation-dependent surround modulation is the plexus of long-range horizontal connections. Recently, Weliky et al., 1995, have shown that in vivo optical imaging and in vitro whole-cell patch-clamp recordings of horizontally evoked excitatory and inhibitory synaptic inputs onto single neurons revealed maximum synaptic inputs when the recording and stimulation electrodes are located within orientation columns sharing the same specificity. In addition, Weliky & Katz (1995) showed that the amplitude of synaptic inputs onto single cells evoked from distant cortical sites is modulated by a cyclical pattern of large and small amplitude responses, with the maximum correlation for neighboring cells and gradual shifts with changes in recording position. Therefore, it can be concluded that the strengths of long-range horizontal connections are maximum between distant cortical cells with the same orientation preference, and they gradually decrease with increasing the relative preferred orientation between distant cells.

I here propose a model of short and long-range cortical dynamics in V1 that helps resolve the apparently conflicting orientation-dependent effects. The model investigates the effect of three major types of input to cortical neurons: feedforward input (from LGN to layer 4 and from layer 4 to layer 2-3), input from intrinsic circuits via horizontal connections (short and long-range corticocortical connections), and input via cortical excitatory feedback (corticocortical feedback from layer 6 to layer 4). The model assumptions are different from those of other theoretical investigations of context effects in V1. These differences are: (a) the use of orientation-dependent connection strengths of long-range horizontal projections, with the strengths decreasing with the increase in the relative orientation between

distant cortical sites; (b) the use of the massive excitatory cortical feedback from layer 6 to layer 4, in addition to local recurrent excitation. The model suggests that the paradoxical dynamic modulation by the surround is due to changes in the gain of the local circuitry as an effect of a disinhibitory mechanism at the level of local circuits triggered by the long-range horizontal connections and amplified by the cortical feedback [previous versions of this model were presented elsewhere (Dragoi, 1995; Dragoi & Wolbarsht, 1995; Dragoi, in press)]. Furthermore, I suggest a link between the orientation-dependent context effects and the perception of geometrical illusions, a demonstration being offered in the case of Zollner illusion.

Materials and methods

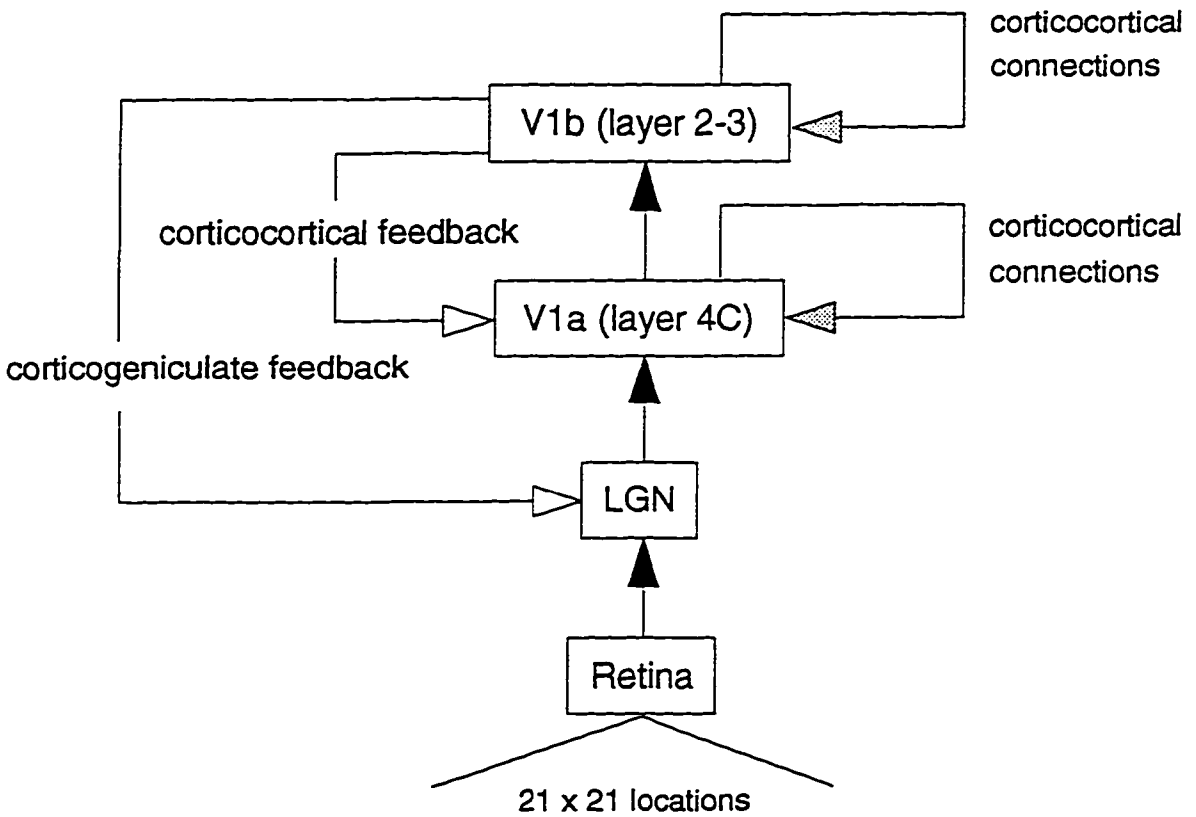
The model

The model simulates parts of the following circuit: retina → LGN → V1a (layer 4C) → V1b (layer 2-3) → V1 (layer 5) → V1 (layer 6) → V1a (Figure 28). For manageability purposes, cells in V1 (layers 5 and 6) are not modeled explicitly; since the projections from layer 2-3 to layer 5 and from layer 5 to layer 6 are purely excitatory the model was simplified by reducing the processing in layers 5 and 6 to a nonlinear filter that integrates the activity of a large number of pyramidal cells in V1b to generate the excitatory cortical feedback to layer 4. Cortical feedback has been hypothesized to amplify the total feedforward input to layer 4 cells by acting as an "active conductance" (Douglas et al., 1995). The nonlinearity is necessary to control the amplification factor on the feedback loop.

For simplicity, a monocular patch of the visual field was divided into 21 x 21 locations, where each location can be uniquely mapped onto a full set of 72 orientation columns (2.5° resolution). The model configures 31,752 LGN neurons arranged on the array of 21 x 21 locations, with 72 cells per each location of the visual patch, and 63,504 cortical neurons arranged on two layers that correspond to V1a (layer 4C) and V1b (layer 2-3).

Figure 28 *Model macrocircuit.* All feedforward pathways (black arrowheads) are excitatory. Corticocortical connections (gray arrowheads) are both excitatory and inhibitory. Corticortical and corticogeniculate feedback pathways (white arrowheads) are excitatory.

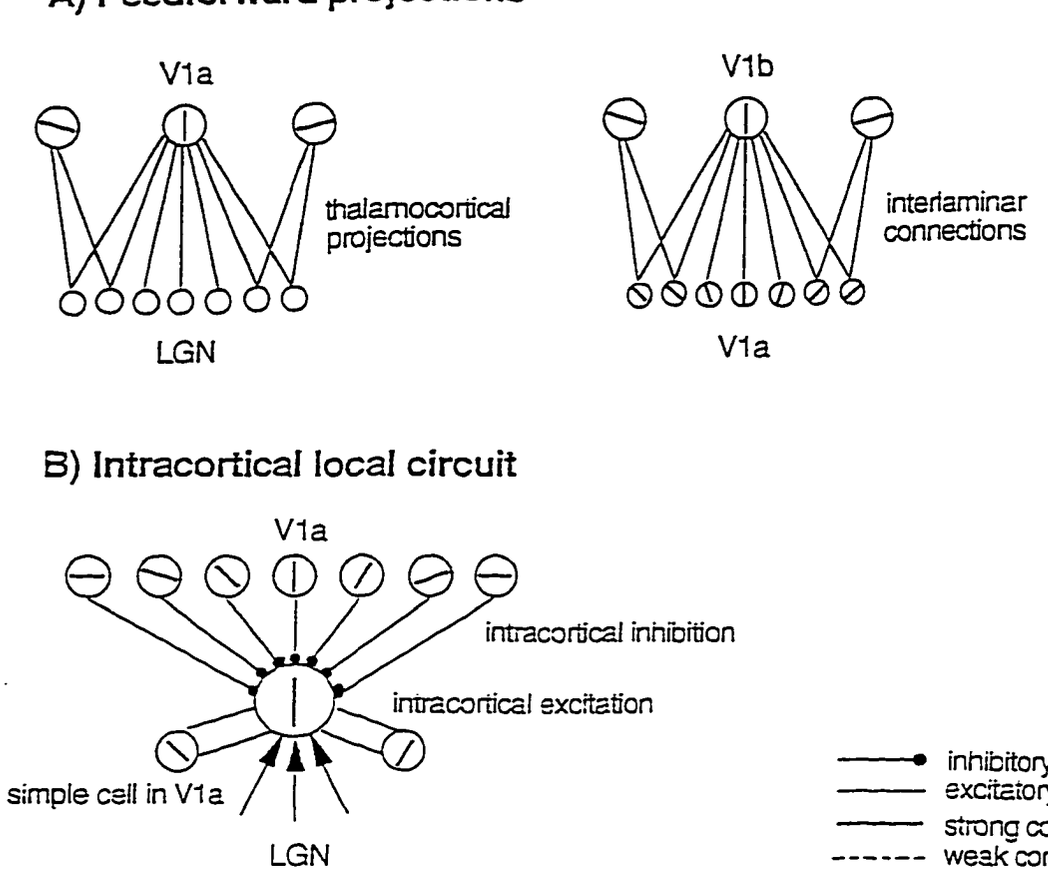
Model



LGN cells receive stimulus-specific input which is maximal for the LGN cell that corresponds topographically to the cortical cell whose orientation preference matches that of the input stimulus [a similar assumption has been used in Stemmler et al. (1995)]. The spread of geniculate inputs to the cortex ensures that each LGN cell synapses within the same hypercolumn to a group of cortical cells within a broad range of orientations (with a spread of 60°). One LGN cell projects to a group of 25 neighboring cortical simple cells in layer 4C centered around the cell which it corresponds topographically. The thalamocortical synapses comprise about 6% of all synapses received by layer 4 neurons (Garey & Powell, 1971; Hornung & Garey, 1981; Winfield & Powell, 1983; LeVay, 1986; Ahmed et al., 1994).

Cortical cells receive center stimulation as an oriented input stimulus applied to the receptive field (RF) center (RF size is idealized to 1 location) and surround stimulation as oriented stimuli applied to the surrounding hypercolumns. Both V1a and V1b cells develop short-range connections within each hypercolumn, i.e., recurrent excitation in the range of 15°, and intracortical inhibition in the range of 60° (Toyama et al. 1981; Michalski et al., 1983; Hata et al., 1988), with the strength of connections decreasing as cortical neurons are more widely separated (Fries et al. 1977; Nelson & Frost 1981), i.e., they differ more in orientation specificity, see Figure 29. Long-range horizontal connections are made only onto cells in V1b within a circular patch of radius 4 at the center. They are made preferentially onto iso-oriented cells (Gilbert & Wiesel, 1989; Malach et al., 1993), the strength of synapses decreasing with the increase in the relative orientation between pre and postsynaptic cells (Weliky & Katz, 1994; Weliky et al., 1995).

Figure 29 *Model connectivity patterns.* (**A**) Feedforward projections from LGN to V1 (layer 4C) and from layer 4 (V1a) to superficial layers (V1b). (**B**) Short-range cortical connections. Both intracortical excitatory (range 15°) and intracortical inhibitory (range 60°) connections are established locally between cells in V1a and V1b. (**C**) Long-range horizontal connections. The strengths of long-range connections are both distance- and orientation-dependent. They decrease with increasing distance and relative orientation between cortical sites (range 45°). (**D**) Cortical feedback. Both corticogeniculate and corticortical feedback projections are excitatory. Both projections are focused in spatial terms and more difuse in the orientation domain.



Model implementation

LGN cells are modeled as single units in which the mean rate of firing, L_i, is given by:

$$\frac{dL_i}{dt} = -0.01L_i + J_iR_i(s + F_i)(1 - L_i) - J_{ij}\sum_{j=i-1}^{i+1} L_jL_i \qquad (11)$$

where the first term, $-0.01L_i$ describes the spontaneous decay in the absence of any stimulation; $J_i = 0.002$ describes the strength of the feedforward connections (one-to-one projections) between retina (R_i) and LGN (L_i); R_i is the stimulus-specific external input to L_i: for an input stimulus of orientation θ, comparable results are obtained using either linear ($R_i = max(0, 1 - \gamma|\theta - \theta_i|)$) or cosine ($R_i = max(0, cos(3\pi(\theta - \theta_i)/36))$ distributions of the input stimulus (computer simulations were run using a linear distribution), where θ_i is the orientation preference of the cortical cell which corresponds topographically with R_i; $J_{ij} = 0.4$ describes the strength of the inhibitory connections between neighboring LGN cells (lateral inhibition); s is a fixed parameter that ensures that the retinal information is processed in the absence of corticogeniculate feedback – introduced for model analysis purposes; F_i is the total feedback that L_i receives from area V1, and it is described by:

$$F_i = h(\sum_j f_{ij}V1b_j) \qquad (12)$$

where $f_{ij} = 0.3$ describes the strength of corticogeniculate connections; $h(x) = \frac{x^{14}}{x^{14}+q^{14}}$ is assumed to be a sigmoidal function ($q = 1.15$); $V1b_j$ are cells in area V1b within a patch of radius 1 (Grieve & Sillito, 1995, have shown that the information carried in the corticogeniculate strem, as well as that carried by axons from layer 6 to layer 4, is relatively tightly focused in spatial terms) that contains projections from a large range of orientation selectivities (Harvey; 1980; Sillito et al., 1993; Grieve & Sillito, 1995), with an orientation spread of 90°). The corticogeniculate synapses outnumber the ascending synapses from LGN to V1a by a factor of 8.

The presynaptic activity of cells in V1b is integrated via the cortical feedback in order to provide global information used to modulate the rate at which the input stimuli are transmitted through the LGN (McClurkin et al., 1994; Mukherjee et al., 1995). This view is supported by recent neurophysiological evidence (Sillito et al., 1995) showing that cortical feedback serves to lock or focus the cortical circuitry onto the stimulus feature by providing global information that can be used to adjust local encoding mechanisms to ensure that information transmission through the LGN is improved.

Cortical cells are modeled as single units in which the mean rate of firing is given by:

- V1a (layer 4C)

$$\frac{dV1a_i}{dt} = -0.01V1a_i + (s+F_i)(J_{ij}\sum_{j=1}^{k1} L_j + J_{ij}^{exc}\sum_{j=1}^{k2} V1a_j)(1-V1a_i) - J_{ij}^{inh}\sum_{j=1}^{k3} V1a_jV1a_i \tag{13}$$

where L_i are geniculate cells that constitute the input for cells in layer 4C (one LGN cell projects within a range of 60°) within the same hypercolumn; $J_{ij} = 1$ is the strength of thalamocortical connections; $J_{ij}^{inh} = 0.2exp^{-0.01|\theta_i-\theta_j|}$ and $J_{ij}^{exc} = 0.1exp^{-0.01|\theta_i-\theta_j|}$ are the orientation-dependent strengths of intracortical inhibitory connections (within a range of 60°) and recurrent excitatory connections (within a range of 15°) from neighboring cells within the same hypercolumn; F_i is the total feedback that $V1a_i$ receives from layer 6; s is a fixed parameter that ensures that the LGN information is processed in the absence of corticocortical feedback (this manipulation is useful when the model is tested in the absence of feedback), and it is described by:

$$F_i = h(\sum_j f_{ij}V1b_j) \tag{14}$$

where $f_{ij} = 0.3$ describes the strength of excitatory cortical feedback connections; $h(x) = \frac{x^{14}}{x^{14}+q^{14}}$ is assumed to be a sigmoidal function ($q = 1.15$); $V1b_j$ are

cells in area V1b within a patch of radius 1 (Grieve & Sillito, 1995, have shown that the information carried by axons from layer 6 to layer 4, is relatively tightly focused in spatial terms); corticortical feedback contains projections from a narrow range of orientation selectivities with an orientation spread of 50° (iso-orientation).

-V1b (layers 2-3)

$$\frac{dV1b_i}{dt} = -0.01V1b_i + J_{ij}\sum_{j=1}^{k1} V1a_j(1-V1b_i) + J_{ij}^{exc}\sum_{j=1}^{k4} V1b_j(1-V1b_i) - J_{ij}^{inh}\sum_{j=1}^{k5} V1b_jV1b_i \quad (15)$$

where $V1a_j$ are cells in V1a that constitute the input for cells in layer 2-3 (one V1a cell projects within a range of $\pm 25^o$ within the same hypercolumn); $J_{ij} = 1$ is the strength of cortical feedforward connections: short-range cortical connections are implemented as described above, when Eq. 14 is introduced, and connection strengths have similar values.

Long-range connections emerge from excitatory cells and contact both inhibitory and excitatory cells (Kisvarday et al., 1986; McGuire et al., 1991). Since at high stimulus contrasts the response of both excitatory and inhibitory cells can be approximated by a linear function (Somers et al., 1995) the long-range effect of one pyramidal cell, $V1b_j$, on another pyramidal cell, $V1b_i$, can be modeled as:

$$V1b_j(J_{ij}^{exc} - dJ_{ij}^{inh}) \quad (16)$$

where $V1b_j$ is the firing rate of the excitatory presynaptic cell; $J_{ij}^{exc} = 0.8exp^{-0.5(\theta_i-\theta_j)\sqrt{r_{ij}}}$ and $J_{ij}^{inh} = 0.4exp^{-0.5(\theta_i-\theta_j)\sqrt{r_{ij}}}$ are the orientation- and distance-dependent synaptic strengths of long-range horizontal excitatory-excitatory and excitatory-inhibitory connections, with θ_i and θ_j being the orientation preferences of $V1b_i$ and $V1b_j$; r_{ij} is the radial distance between $V1b_i$ and $V1b_j$; d is a contrast-dependent dominance factor such that if $d < J_{ij}^{exc}/J_{ij}^{inh}$ the net effect of long-range connections is excitatory, whereas if $d > J_{ij}^{exc}/J_{ij}^{inh}$ the net effect is inhibitory. Since all the stimuli utilized in the original experiments simulated here are presented at high-contrast, and well-documented studies (Hirsch & Gilbert, 1991; Weliky et al., 1995; Toth et al., 1996) have revealed that the effect

of horizontal connections on the target cell is excitatory when the cell receives weak input (low contrast), and inhibitory when the cell receives strong input (high contrast), we set $d = 4$ such that to bias the balance of long-range connections toward a net inhibitory effect.

Results

The key orientation-dependent context effects depend on the properties of the local cortical circuitry, therefore I will explain first the underlying physiological mechanism in qualitative terms, followed by a detailed presentation of the simulation results. Figure 30 describes a model cortical circuit incorporating long- and short-range horizontal connections. Long-range excitatory projections originate from pyramidal cells in the surround and contact other pyramidal cells as well as nearby inhibitory cells, which are locally interconnected (CRF). These connections link preferentially iso-orientation domains; more generally, the strength of synaptic connections between distant cortical sites varies gradually from zero to a maximum depending on the difference in orientation preference (Gilbert & Wiesel, 1989; Malach et al., 1993; Weliky & Katz, 1994; Weliky et al., 1995). Therefore, activation of horizontal connections can evoke direct iso-orientation excitatory and multisynaptic inhibitory inputs onto local pyramidal cells in an orientation-dependent fashion: stronger activation of iso-orientation domains and gradually weaker activation of non-iso-orientation domains. Since the context effects to be explained in this study have been obtained at high contrast for both classical and nonclassical receptive field stimulations, and such high contrast stimuli were shown previously to bias the relative balance of excitatory and inhibitory inputs toward strong inhibition (Weliky et al., 1995; Toth et al., 1996), there will be stronger activation of inhibitory cells in the nearby of iso-orientation domains and weaker activation of inhibitory cells in the nearby of non-iso-orientation domains. An example is shown in Figure 30 in which the "iso-orientation" inhibitory cell on the left is activated more strongly than the "non-iso-orientation" inhibitory cell on the right.

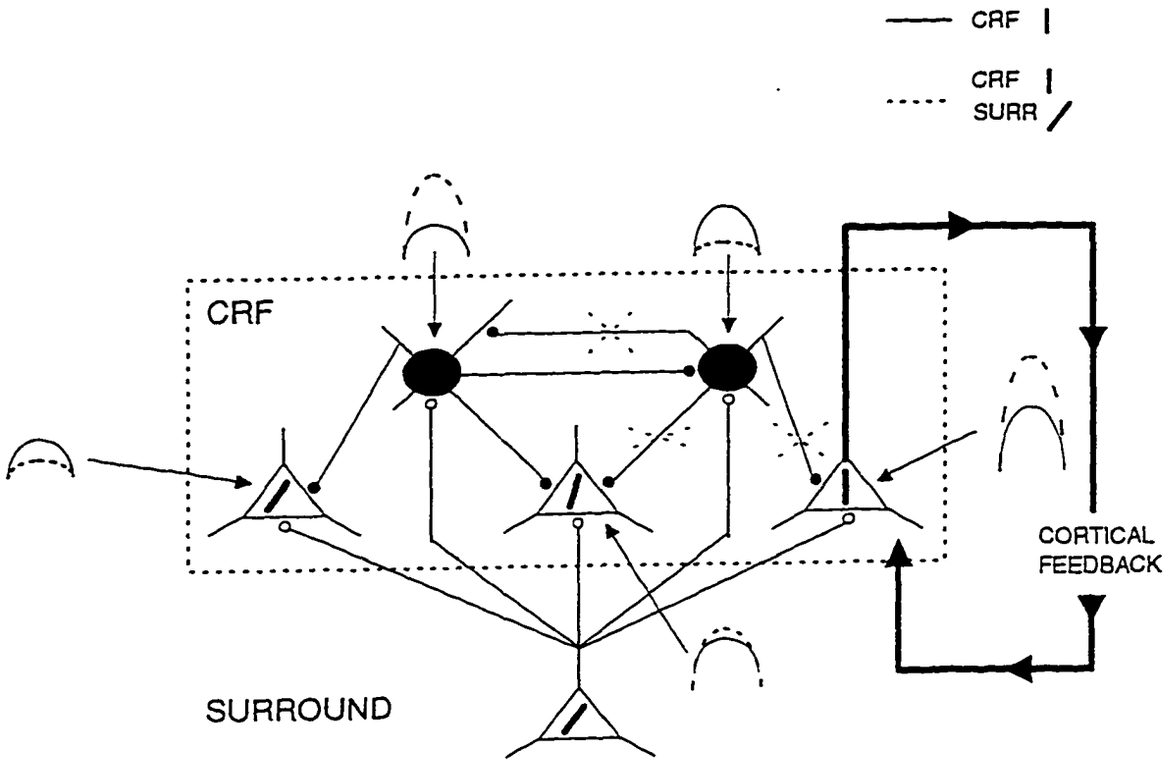
Because inhibitory cells are well-known for their lack of adaptation (McCormick et al., 1985), "iso-oriented" inhibitory cells are able to fire continuously in response to a tonic center and surround stimulus and thus exert a tonic inhibition on their postsynaptic targets, including other inhibitory cells. The effect of this interaction is the disinhibition of "non-iso-orientation" inhibitory neurons (e.g., see Figure 30 – the inhibitory cell on the right). This disinhibitory interaction can remove tonic inhibition from pyramidal cells in the nearby of "non-iso-orientation" inhibitory neurons (e.g., see Figure 30 – the pyramidal cell that prefers vertical orientation). This type of interaction has been previously observed in the pericruciate cortex of the cat (Kelly & Renaud, 1974). Thus, the net effect is that upon the application of an oriented surround in conjunction with a center stimulus oriented away from the surround, non-iso-oriented pyramidal cells are released from local inhibition and fire more vigorously than under no-surround condition (at the same optimum center stimulus).

Context-dependent removal of inhibition through local disinhibition is an intricate process that yields an orientation-dependent dynamic gain control mechanism: an oriented surround increases the responsiveness of non-iso-oriented pyramidal cells, which respond tonically to the thalamocortical input, via a disinhibitory mechanism, while iso-oriented pyramidal cells are strongly suppressed.

The orientation-dependent facilitatory (disinhibitory) effect of the surround can be further amplified by the recurrent excitatory cortical feedback connections from layer 6, and for some center-surround configurations the combined effect of disinhibition and feedback amplification can bring pyramidal cells to high "supra-optimal" response levels (e.g., Sillito et al., 1995).

Large scale computer simulations were performed to evaluate neuronal responses to oriented stimuli that covered (a) the classical receptive field (CRF) alone (center stimulus); (b) the classical and extraclassical receptive fields (center + surround stimuli).

Figure 30 *Schematic diagram of the interactions between representative cells in V1b embedded in their local network.* Cells in the surround project on both excitatory (pyramids) and inhibitory (filled circles) neurons. Since the strength of long-range projections is orientation-dependent, there will be an increased activation of the "iso-orientation" inhibitory neuron on the left which suppresses the "non-iso-orientation" inhibitory neuron on the right. Therefore, cells which prefer stimuli oriented away from the surround orientation will be released from inhibition (disinhibition). This effect is further amplified by the excitatory cortical feedback which bring responses to a supra-optimal level. Solid line – activity patterns when the CRF is stimulated with a vertical bar; Dashed line – activity patterns when the CRF is stimulated with a vertical bar while the surround is stimulated with a tilted bar. Empty circles – excitatory connections; Filled circles – inhibitory connections.



Center and surround stimuli were always presented at high contrast (fixed value in the simulations). Both center and surround orientations were varied systematically from 0 to 180° to fully investigate the orientation-dependency of context effects. The surround stimulus modulates the response to the center in an orientation-dependent fashion. Responses to the center stimulus are suppressed by an iso-oriented surround, whereas responses to the same center stimulus become "supra-optimal" in the presence of an orthogonal or oblique surround. The model predicts that orientation-dependent strengths of long- and short-range connections can have bi-phasic modulatory effects, depending on the relative orientation between center and surround. To illustrate the model's key properties computer simulations will be presented to analyze surround facilitatory effects (Knierim & Van Essen, 1992; Sillito et al., 1995) and surround suppressive effects (Knierim & Van Essen, 1992). Related experimental results (e.g., Grinvald et al., 1994) will be discussed without showing computer simulations. In addition, the same mechanism will be applied to demonstrate that the model is able to explain the surround-dependent orientation shift effect (Gilbert & Wiesel, 1990). Finally, I will show that the same mechanism responsible for physiological orientation-dependent context effects underlies psychophysical effects, such as geometrical illusions of orientation (an example is the Hering illusion).

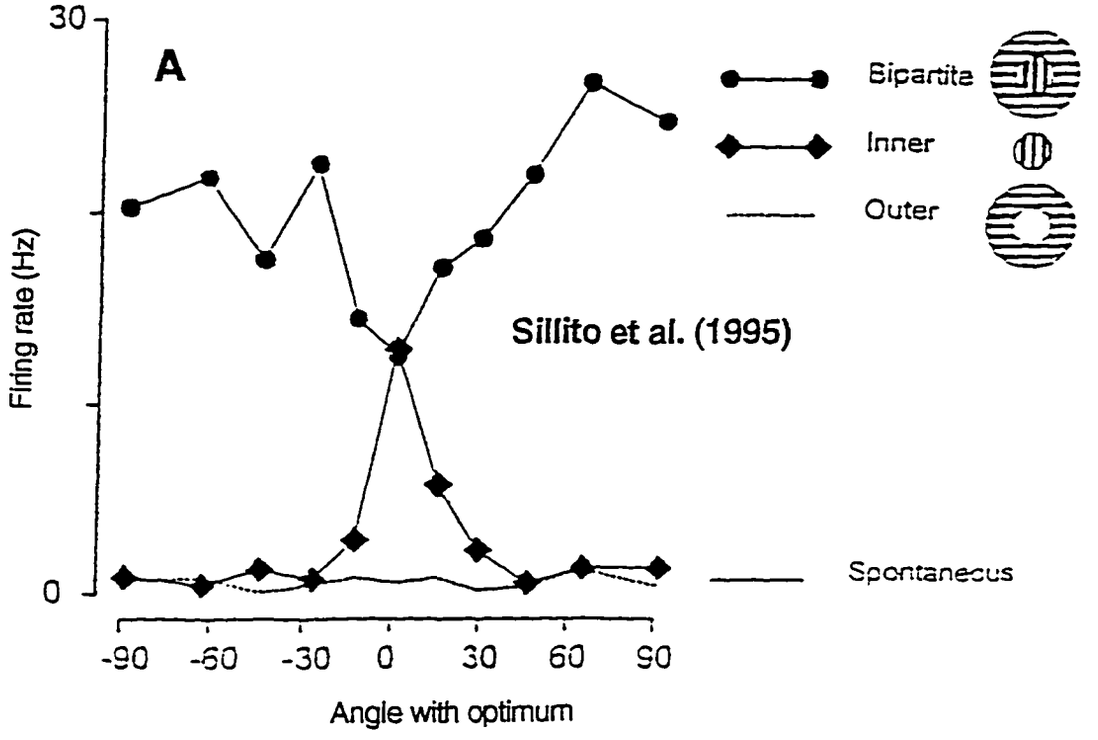
Orientation-dependent surround suppression and facilitation

Representative experimental data showing both suppressive and facilitatory surround effects are presented in Figures 31A and 32A-B. If the RF is stimulated with an oriented bar, iso-oriented stimuli in the surround cause suppressive effects, while cross-oriented stimuli in the surround cause supra-optimal responses in simple cells in primate V1. In Sillito et al.'s (1995) experiment (Figure 31A) the surround (annulus) alone has no effect on the cell's response, which in this case fluctuates around the spontaneous level of firing. However, as the surround orientation is varied while the center (inner) stimulus is presented at the cell's preferred

orientation, the surround strongly modifies the response. For surround orientations between 30° and 90° relative to the cell's optimal orientation the response is strongly increased. Simulation results are shown in Figure 31B. For iso-oriented surround stimuli the response to the center is strongly suppressed (although Figure 31A does not reflect the iso-orientation suppression, other figures in Sillito et al.'s (1995) study such as Figure 1, panels b, c, d, show strong iso-orientation suppressive effects). The model explains iso-orientation suppression by the fact that iso-oriented center and surround stimuli activate maximally local "iso-orientation" inhibitory cells that suppress nearby pyramidal cells. In addition, because of the spread of intracortical inhibition (60° in the model) suppressive effects also exist when the relative orientation between center and surround increases (e.g., 30° in Figure 31B), although not sufficiently to trigger facilitatory effects.

If the surround orientation is varied further away from the cell's preferred orientation (which is 90°), at some critical difference between center and surround orientations, equal to the spread of long-range horizontal connections (45°), the surround is no longer able to directly influence the "vertical" cells. However, the surround is still able to exert indirect influences on "vertical" cells through the disinhibitory mechanism described in the previous section. Therefore, if the orientation mismatch between center and surround is greater than 45° the cell's response becomes supra-optimal (weak surround inhibition combined with local disinhibition). Furthermore, as the surround orientation deviates further from 45° the strength of local inhibition gradually diminishes increasing the gain of the local circuit, which now favors strong excitation; the cell's response becomes more supra-optimal. However, if the relative orientation between surround and center increases beyond 60°, which is equal to the spread of intracortical inhibitory connections in the model, the strength of the disinhibitory mechanism diminishes because "iso-orientation" inhibitory cells do not extend so far and their contribution to the disinhibitory mechanism gradually vanishes.

Figure 31 *Orientation-dependent surround facilitation.* (**A**) Supra-optimal response in a simple cell in primate V1 (Sillito et al., 1995). As the orientation of the annulus is varied with the inner patch present at the optimum orientation, for orientations of the outer between 30° and 90° the cell's response is strongly increased. (**B**) Supra-optimal response in cells in V1 (theoretical tuning curves).



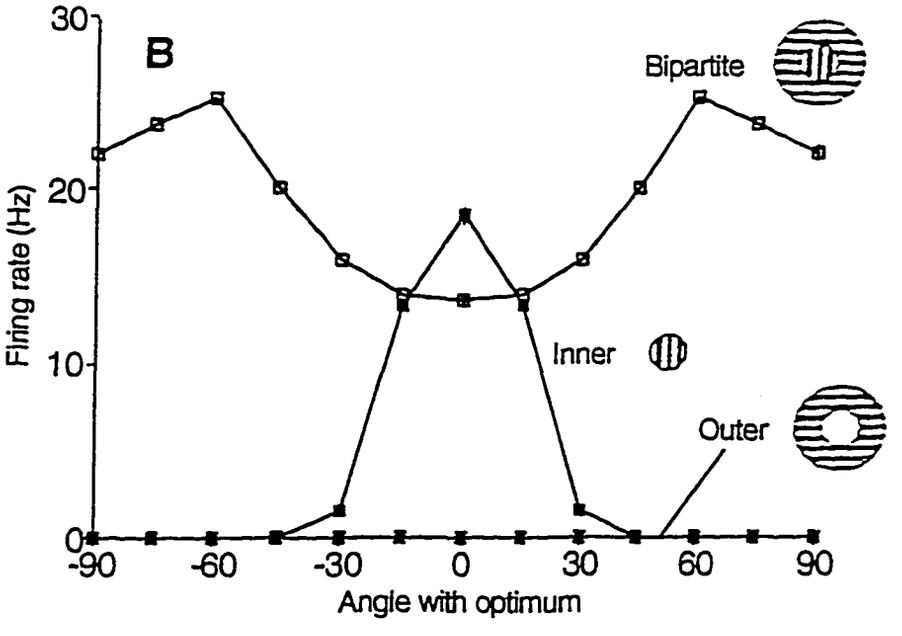


Figure 32A shows normalized population responses in V1 (n=50) when cells are stimulated at their optimum orientation (vertical) while the surround orientation pattern is varied (Knierim & Van Essen, 1992). Figure 32B shows the result of similar surround manipulations, except that normalized individual responses are shown (therefore it is not surprising that the normalized response magnitudes from panels A and B differ, mainly with respect to the response to the orthogonal surround). According to the model, Figure 32C, maximum suppression is obtained when the surround is parallel with the center, and, as the surround orientation becomes random, its inhibitory influence diminishes because the strength of long-range connections decreases with the mismatch in orientation between center and surround. When the surround and the center are orthogonal, the disinhibitory mechanism described in the previous section is engaged and the cell's response becomes supra-optimal. This stimulus situation favors psychophysically the perceptual effect of pop-out, i.e., facilitation in the detection of features which are very different from context. Pop-out is correlated with supra-optimal responses in cells "attending" the target stimulus. When the center stimulus is absent, the cell whose response rate is measured shows only the background level of activity.

Orientation shift effect

One important instance of surround modulation is the context-dependent orientation shift effect (Gilbert & Wiesel, 1990) – if a cell in the superficial layers of V1, that shows preference for 30° stimuli, is presented a surround oriented 30° counterclockwise from the optimal orientation, the cell's tuning curve is shifted 10° in a direction away from the orientation of the surround (Figure 33A). Figure 33B shows tuning curves generated when the model is exposed to the same conditions as in the original experiment. The model explains the orientation shift effect using the disinhibitory influence exerted by the long-range horizontal connections that release the orientation detectors from local intracortical inhibition along a direction away from the surround orientation.

Figure 32 *Orientation-dependent surround suppression and facilitation.* (**A**) Surround suppressive effect (normalized responses for cortical cells tested with different orientation stimuli). The maximum suppression is obtained at iso-orientation, the strength of suppression reducing as the difference between center and surround orientation approaches 90^o. This figure was adapted after Figure 11 of Knierim and Van Essen (1992). (**B**) Normalized individual responses showing an orientation contrast effect (supra-optimal responses) when center and surround stimuli are orthogonal. (**C**) Surround suppressive effect (theoretical normalized responses). This figure should be compared with the experimental results in Figure 32A.

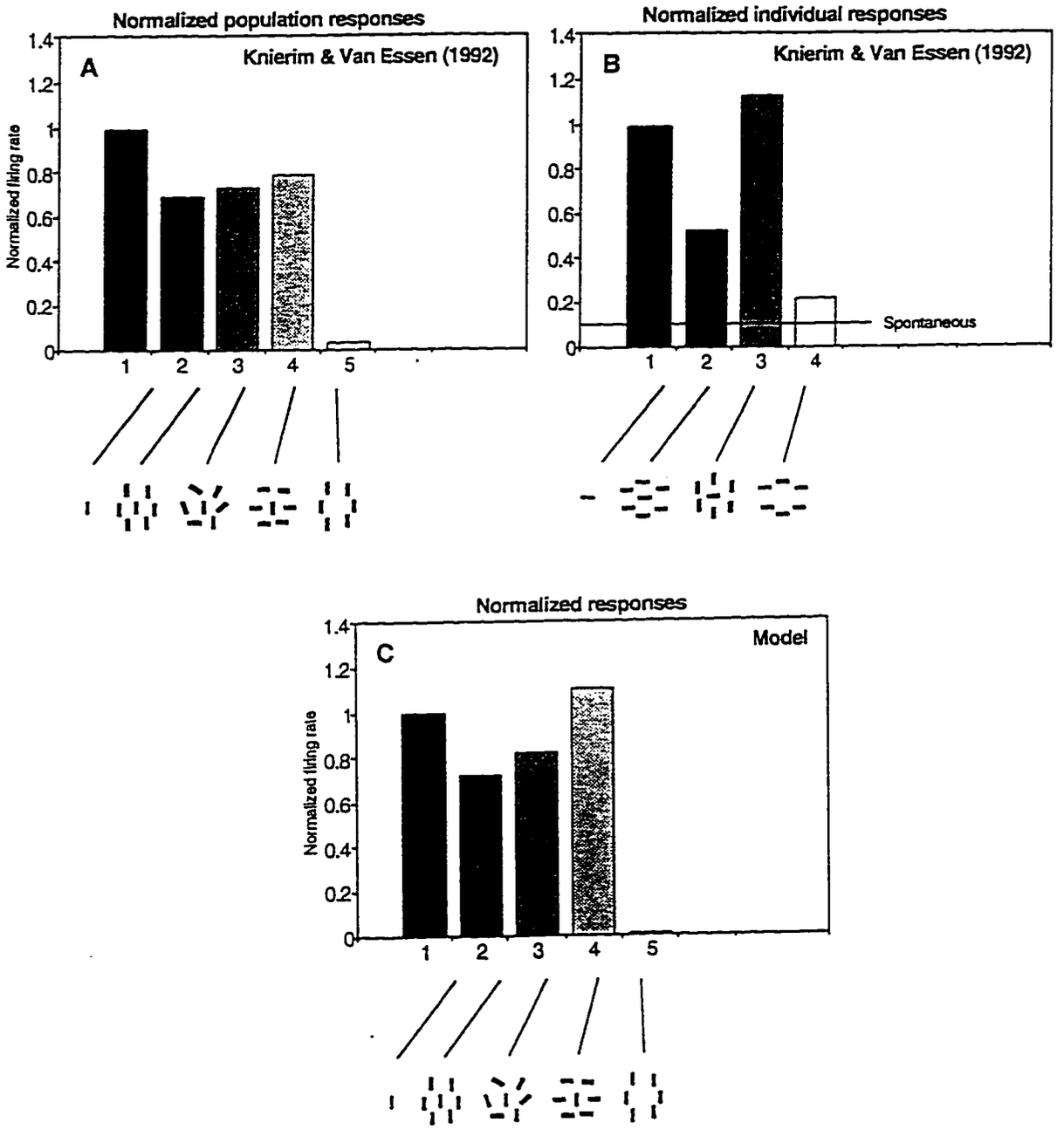
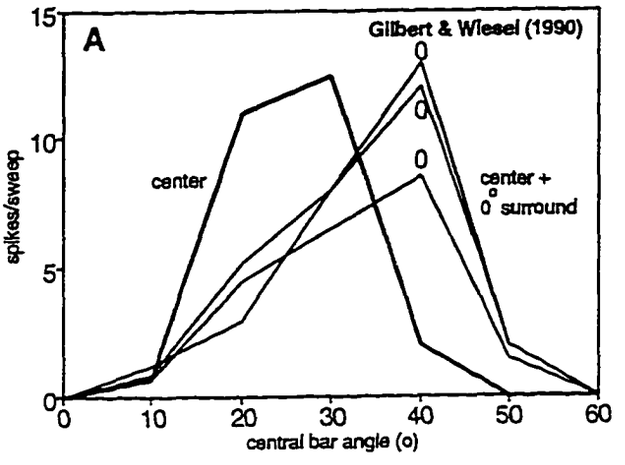
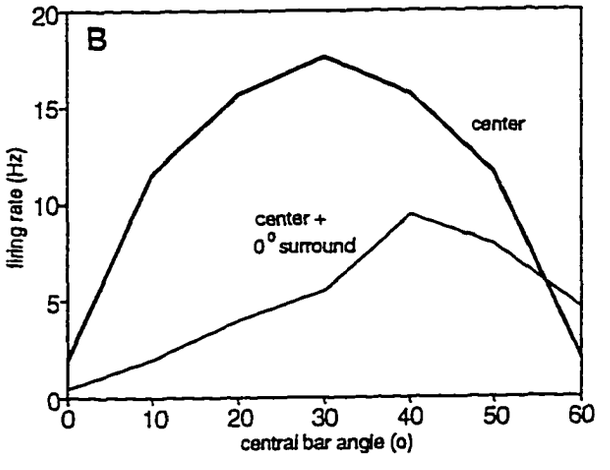
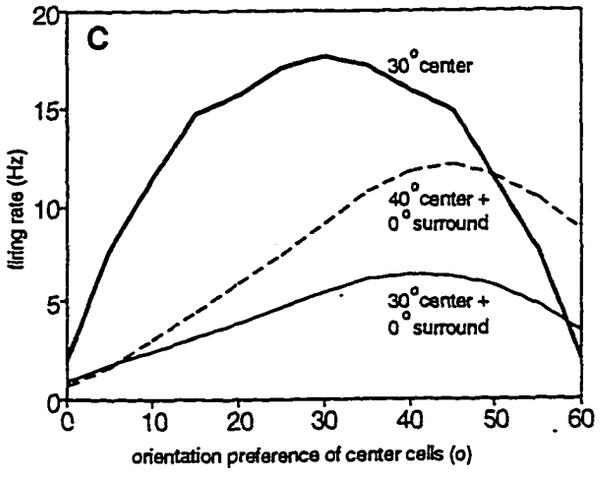
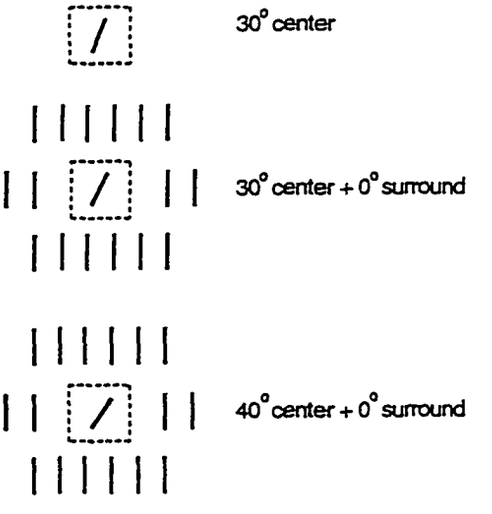


Figure 33 *Context-dependent orientation shift effect.* (**A**) Orientation tuning curves for a superficial layer cell without surround (bold line) and with a 0^o surround (simple line). Even though the cell's initial tuning curve peaked at 30^o, each time the surround is present there is a shift in the cell's orientation specificity to 40^o in a direction away from the surround. This figure was adapted after Figure 4 of Gilbert and Wiesel (1990). (**B**) Context-dependent orientation shift effect (theoretical tuning curves). This figure should be compared with the experimental results in Figure 33A. (**C**) Cortical response histogram. Right – stimulus arrangement in three experimental conditions: 30^o stimulus in the CRF; 30^o stimulus in the CRF in conjunction with a 0^o stimulus in the surround; 40^o stimulus in the CRF in conjunction with a 0^o stimulus in the surround.







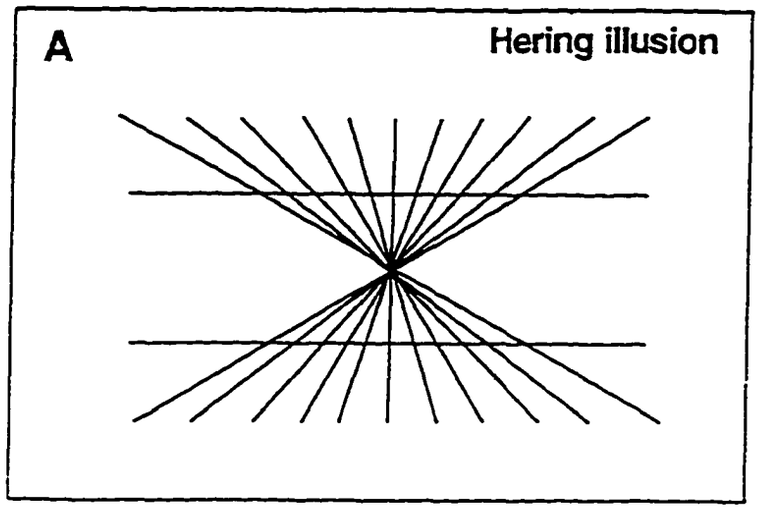


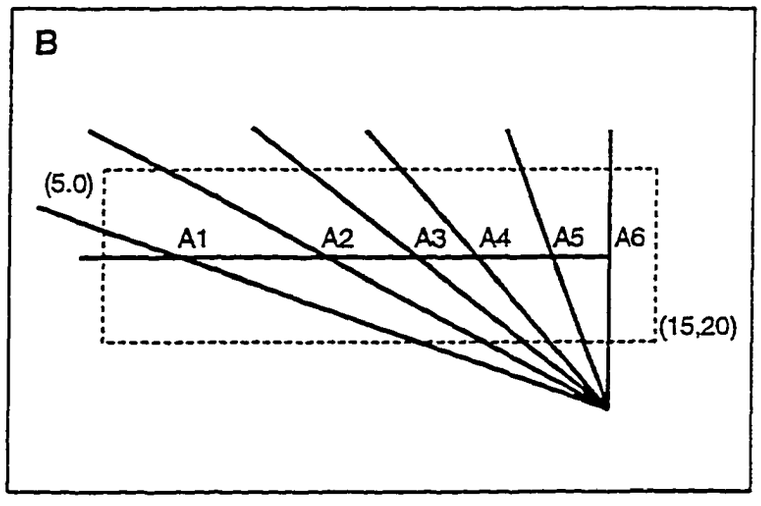
When the 0° surround is applied, the net contribution of long-range connections is strong suppression of cells with orientation specificities nearer the surround (e.g., cells tuned to 30°). Therefore, the presynaptic potential with which these cells control the strength of local inhibition to neighboring cells within the same hypercolumn diminishes. Under these conditions, cells in the center which prefer stimuli (e.g., 40°) oriented away in a direction clockwise from surround orientation and are only weakly suppressed by the surround (Figure 33C). are released from intracortical inhibition and respond to the excitatory feedforward input received via the spread of the thalamo-cortical afferents. If the surround is maintained fixed (0°) while the center is stimulated with a bar oriented at 40°, the overall responsiveness of the cells which receive center stimulation increases relative to the situation when the center stimulus is oriented nearer the surround (30°), such that the center stimulus is able to elicit a higher response in the cell whose original tuning curve peaked at 30° (Figure 33C). This orientation-dependent increase in cortical responsiveness is further amplified by the cortical feedback. The repulsive shift in orientation specificity is thereby a byproduct of the interplay between surround inhibition, intracortical inhibition and cortical feedback.

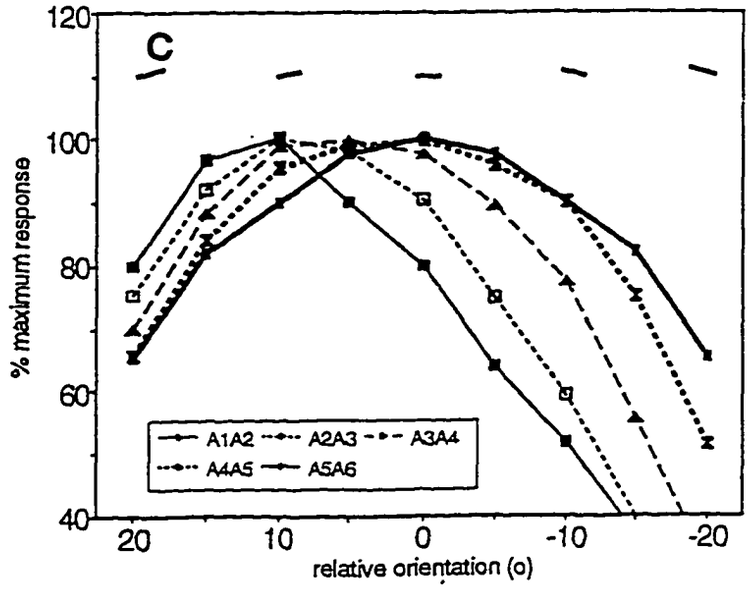
Geometrical illusions of orientation (Hering)

To test how the model performs in more complex situations resembling geometrical illusions (see Robinson, 1972, and Gillam, 1990, for a review), the upper left configuration of the Hering pattern (Figure 34A) was projected onto a window of 21 x 10 locations (Figure 34B) of the network. The horizontal line of the Hering configuration is divided by the obliques into six short line segments. Viewed from left to the right, each one of these horizontal short lines receives orientation surrounds that gradually approach 90°. According to the previously discussed context-dependent orientation shift effect, the magnitude of the shift is inversely proportional to the orientation mismatch between center and surround.

Figure 34 *Geometrical illusions of orientation.* (**A**) Hering geometrical illusion. The geometrical arrangement of the oblique lines (surround) in the Hering configuration creates conditions for the two horizontal lines to be perceived as curved on the center. (**B**) The upper left Hering configuration is projected onto a window of the location map (dotted rectangle). Each of the short horizontal bars between the obliques has a size of 3 locations, excepting A_1A_2 which has the size of 4 locations; the horizontal bar from outside the oblique bars has a size of 3 locations. Viewed from left to the right, the orientations of the obliques are: 20°, 30°, 40°, 50°, 70°, and 90°. (**C**) Percentage maximum response histograms of cells receiving A_1A_2, A_2A_3, A_3A_4, A_4A_5, and A_5A_6 as center stimulus. The x-axis represents the difference between the orientation specificity of different cells within the same hypercolumn and the 0° orientation (for instance, a relative orientation of 10° represents a horizontal bar tilted 10° counterclockwise). The histograms have been obtained by averaging the percentage maximum responses of cells of the same orientation preference in the three adjacent hypercolumns responding to A_1A_2, A_2A_3, A_3A_4, A_4A_5, and A_5A_6 respectively. Different cells show maximum responses that qualitatively match the perceived orientation of each of the 5 line segments in the Hering configuration.







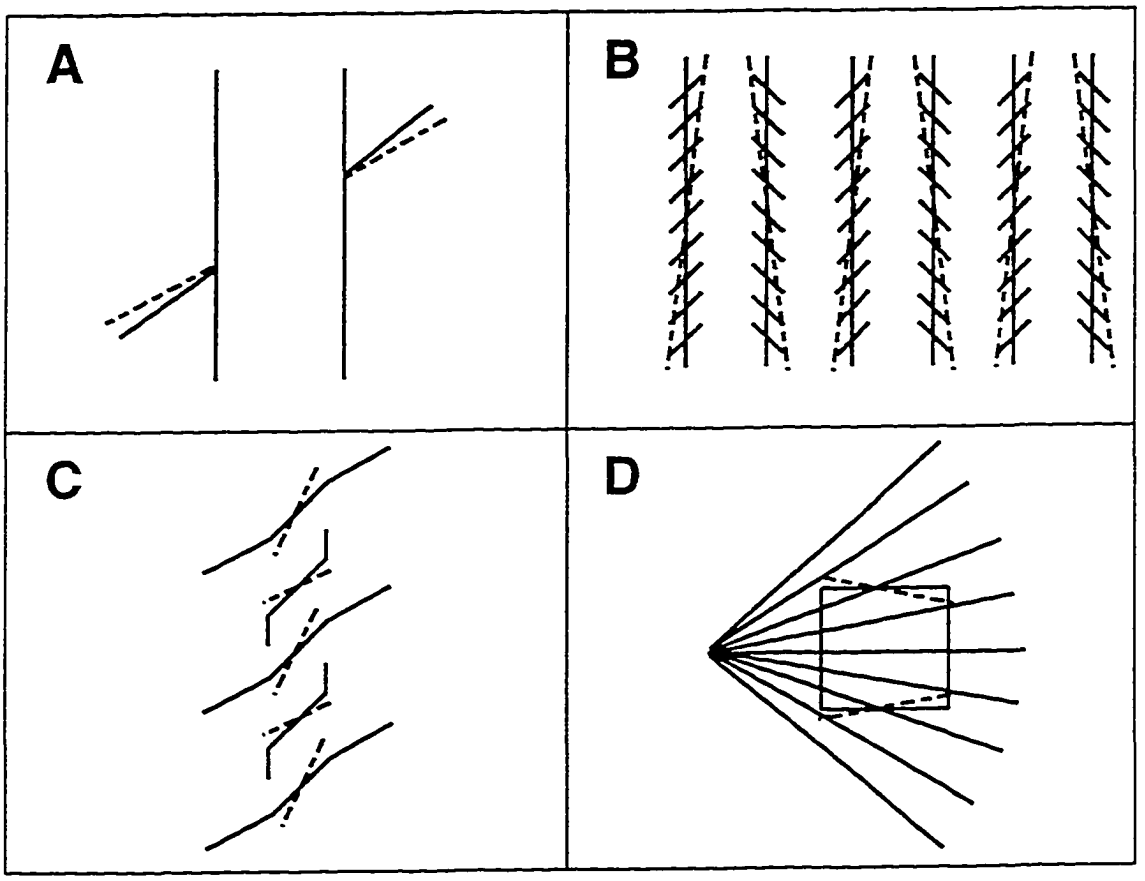
In the case of the Hering pattern (Figure 34A), there is a greater shift (in a direction away from the surround) in orientation preference for the cells responding to the line segment A_1A_2 than for the cells responding to the line segment A_2A_3. Also, there is a greater shift in orientation preference for the cells responding to the line segment A_2A_3 than for the cells responding to A_3A_4, and only a negligible shift for the cells responding to A_4A_5 and A_5A_6 (the surround approaches 90°). Simulation results (Figure 34C) show groups of cells peaking at orientations that match the illusory percept.

In a similar way, the model can connect the surround effects with a large number of geometrical illusions, among the most representative being Poggendorff, Zollner, Lipps, Ehrenstein (Figure 35).

Discussion

This article proposes a model of orientation-dependent context effects that relies on a dynamic gain control mechanism controlled by the interplay between orientation-dependent long-range connections, intracortical excitatory and inhibitory connections, and cortical feedback. The model accounts for orientation-dependent surround suppression and facilitation demonstrated in several studies, as well as for context-dependent orientation shift effects. Previous theoretical studies (Somers et al., 1995; Stemmler et al., 1995) have suggested that stimulus contrast is an independent variable that needs to be considered when the properties of classical and non-classical receptive field are evaluated. The present study focuses exclusively on stimulus orientation as an independent variable that controls whether the surround role is either suppressive or facilitatory, depending on the relative orientation between center and surround stimuli. The model has the potential to explain the emergence of excitatory responses to cross-oriented stimuli presented in the classical receptive field in conjunction with iso-oriented surround stimulation (Sillito et al., 1995).

Figure 35 *Selected geometrical illusions that can be explained by the model.* (**A**) Poggendorff - the oblique lines are collinear. (**B**) Zollner - the vertical lines are parallel. (**C**) Lipps - the oblique lines in the middle are parallel. (**D**) Ehrenstein - the figure in the center is a square.



The iso-oriented surround strongly suppresses responses to an iso-oriented center stimulus; however, when an orthogonal stimulus is presented in the classical receptive field the "cross-oriented" cells receive their preferred stimulus in the center and are disinhibited from local inhibition through the dynamic gain control mechanism described previously. Therefore their response becomes supra-optimal and via the orientation spread of the cortical feedback from layer 6 (perhaps combined with corticogeniculate feedback from layer 6 to LGN) the response of "iso-oriented" cells increases. This view suggests a new role for cortical feedback (corticocortical and/or corticogeniculate projections), that is cortical feedback may have important implications with respect to the capacity of the feedback projections to carry information about the visual context to the geniculate and layer 4C level (it is known that LGN and layer 4 lack the anatomical substrate that underlies the surround modulation). In fact, Sillito et al. (personal communication) found both suppressive and supra-optimal responses in both LGN and simple cells in layer 4. This type of interaction is impossible to explain if we rely *only* on long-range connections, without involving some form of feedback from the superficial layers of V1, or maybe cortical feedback from extrastriate areas.

Besides explaining several physiological phenomena, the orientation-dependent model presented here demonstrates that the same key principles can be used to investigate the nature of orientation-dependent psychophysical distortions of lines such as geometrical illusions of orientation (a demonstration is offered for the Hering illusion). It can also be speculated that supra-optimal responses in V1 induced by cross-oriented surround stimuli underlie the perceptual "pop-out", and they may also play a role in the figure-ground discrimination and texture segregation.

In addition to offering a new view on the orientation-dependent long-range interactions in primary visual cortex by proposing a testable dynamic gain control mechanism at the level of local circuitry, the model generates other predictions: (a) if stimulus contrast, in addition to orientation, is varied, for a fixed relative orientation between center and surround which generates the maximum effect in Sillito

et al.'s (1995) experiment, e.g., surround tilted 60° relative to the cell's preferred orientation, a low-contrast surround stimulus is able to abolish the supra-optimal response, even though the center is stimulated at low-contrast (it is known that low-contrast stimuli presented in the classical receptive field facilitate responses in the presence of a high-contrast surround); (b) experiments performed in V1 of primates would reveal patterns of response similar to those shown in Figure 34B in relation to the Hering illusion. In addition, the model predicts that the tuning curve of cells with receptive fields covering the horizontal line gradually shift their preference in the presence of the oblique surround such that the shift magnitude should decrease as the surround gradually approaches 90°.

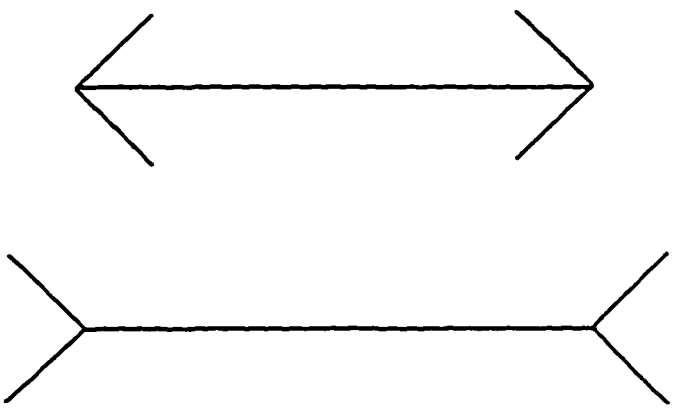
Context-dependent changes in visual sensitivity and the Müller-Lyer illusion

Abstract

Spatial vision is context dependent, i.e., the perceived visual attributes of a target stimulus depend on the context within which the target is placed. Geometrical illusions, which are context-induced subjective distortions of visual features, such as length, orientation, or curvature of lines, are the most striking example. Perhaps one of the best-known and most extensively investigated geometrical illusions, the Müller-Lyer configuration (Müller-Lyer, 1889 : see Figure 36), in which a line appears short or long when it is flanked by outward or inward-pointing arrowheads, has fascinated researchers for over 100 years. Most proposed theories of the Müller-Lyer illusion rely on 'later' processes, hypothesized to occur after 'early' processing of the visual image by local mechanisms selective for features such as orientation and spatial frequency. These theories assume some sort of confusion, or blurring of features, that occurs when observers judge the length of the horizontal bar: confusion of the embedded depth information (Gregory, 1963; Fisher, 1967), confusion of the distance between arrowhead tips (Pressey, 1970; Erlebacher & Sekuler, 1969), confusion of the location of the arrowhead vertex (Chiang, 1968), or confusion during binding of contour fragments (Rensink & Enns, 1995). One problem with these theories of Müller-Lyer illusion is that they leave the concept of 'confusion' largely undefined. In contrast, here I attempt to disambiguate this concept showing striking correlations between the perceived length distortion in the Müller-Lyer illusion and 'low level' visual processing such as detectability of a luminance bar

(target), in stimulus configurations that contain a single set of inward and outward-pointing arrowheads. I found that the inward-pointing arrowhead improves target detectability, whereas the outward-pointing arrowhead suppresses target detectability. Both effects diminish as target contrast, arrowhead angle, and target-arrowhead spatial disparity are increased. At larger distances between target and arrowhead the suppressive effects are completely abolished. The contrast detection results reported here suggest a common mechanism for target detectability and perceived extent. I explore the properties of a population model of orientation detectors in visual cortex that relies on short and long-range horizontal cortical connections to explain the nature of both context-induced suppressive and facilitatory effects in contrast detection.

Figure 36 *Müller-Lyer geometrical illusion.* The horizontal bar appears longer in the inward-pointing-arrowhead condition than in the outward-pointing-arrowhead condition.



Introduction

Although different explanations for the occurrence of the Müller-Lyer illusion have been advanced, e.g., depth theories (Gregory, 1963; Fischer, 1967 – length distortions are due to misapplication or confusion of size constancy to the two spans), averaging theories (Erlebacher & Sekuler, 1969; Pressey, 1970 – the arrowheads interfere with the perceptual system for measuring span of the horizontals, and therefore observers confuse or average the distance between the arrowhead tips), displaced vertex theories (Chiang, 1968 – the perceptual system miscalculates the location of the arrowhead vertex, displacing it toward the concave side), they do not agree on the basic perceptual principles that underlie the apparent distortions.

The present study analyzes the influence on the detectability of a target stimulus of visual contexts which resemble the Müller-Lyer illusion (the perceived distance between two arrowheads is greater as the angle size increases, and this effect diminishes in magnitude as the arrowheads and line are separate more and more – Robinson, 1972). I show striking correlations between the Müller-Lyer extent illusion and target detectability in the same stimulus configurations. For simplicity, the first study used a configuration with only one set of fins (two segments that join to form a vertex), rather than two, to examine the detectability of a target stimulus (this configuration generates the extent effect originally reported with the Müller-Lyer configuration – Greene & Nelson, 1997).

Psychophysical tests

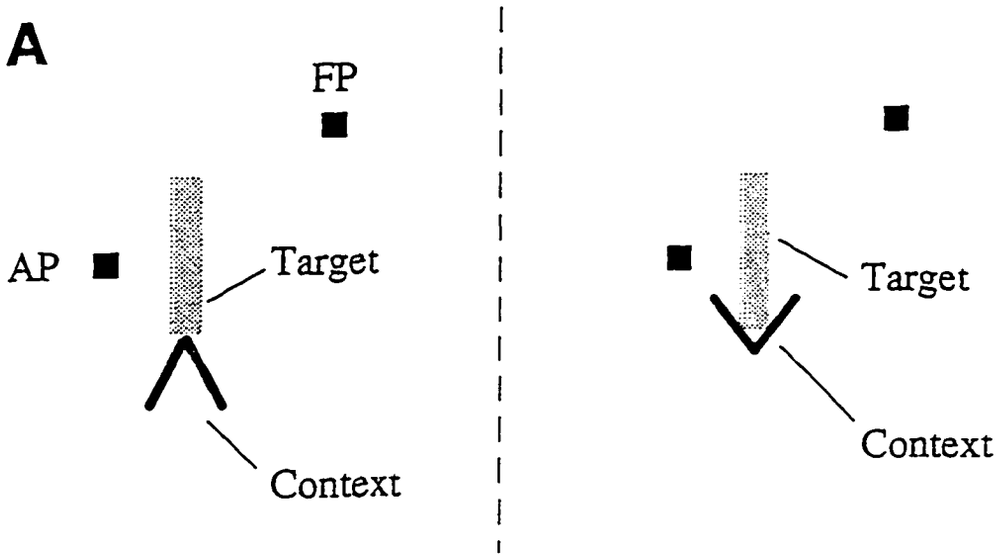
In individually conducted sessions, six adult human observers were instructed to detect the occurrence of a subthreshold vertical bar (target) while the target and a high-contrast arrowhead (context) were briefly flashed periodically. Figure 37A shows a schematic representation of stimulus configuration. A 8.5′ square fixation point (FP) and a similar square attention point (AP) always remained on the screen. Subjects were instructed to fixate at FP and to attend to AP. At an

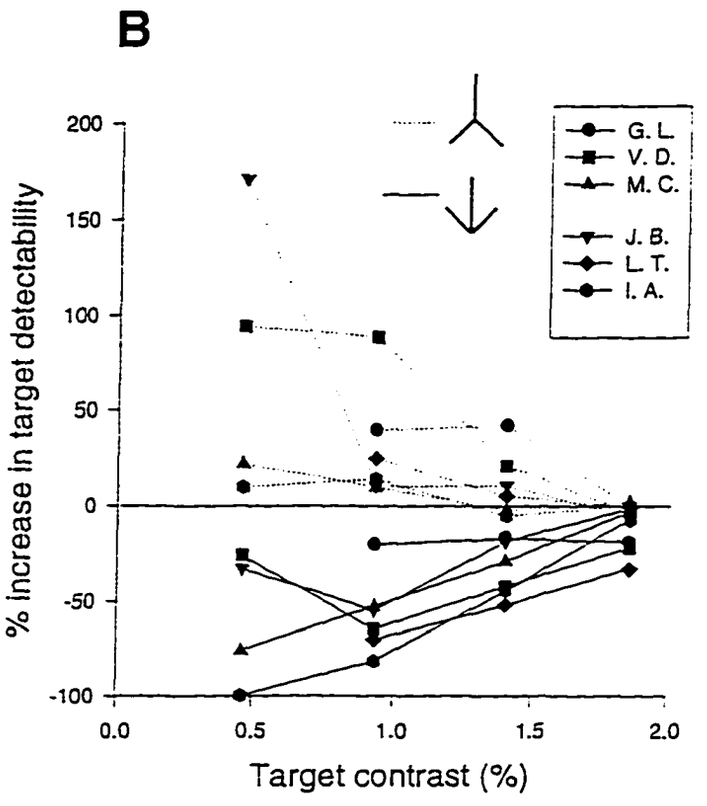
observation distance of 100 cm the target consisted of a vertical line (1^o x $36'$) against a uniform background of 35.71 cd/m^2. The target was always presented $20'$ of arc to the right of AP and 3^o downward and to the left of FP. Each trial consisted of a 2000 ms cycle. The target and/or the context were presented for 200 ms followed by 1800 ms interstimulus interval. The target occurred randomly at one of four contrasts (0.46%, 0.93%, 1.4%, and 1.86%); the context (outward or inward arrowhead with the fin length of $52'$ x 14') was always presented at the same high contrast (0.17 cd/m^2) as FP and AP. During each trial the target was presented alone, in conjunction with the context, or neither target nor context were presented. Whenever subjects detected the target they were instructed to press the mouse button. I measured the detectability ratio, defined as proportion correct detections, at four target contrasts in the following three conditions: (a) target alone; (b) target and outward-pointing arrowhead; (c) target and inward-pointing arrowhead.

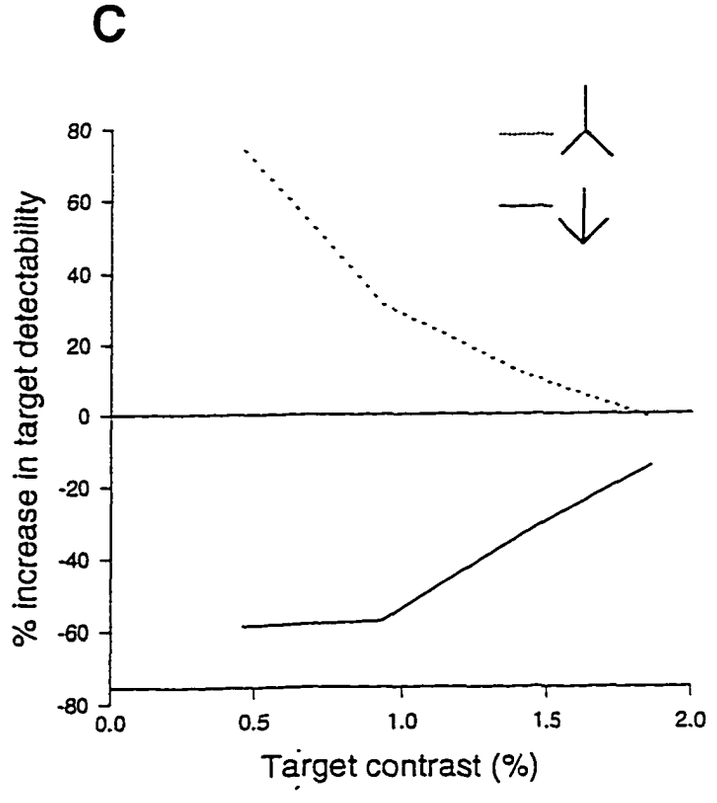
Compared to when no target was shown, I found that for each individual subject target detection is facilitated by the inward arrowhead and suppressed by the outward arrowhead (Figure 37B). Furthermore, again for each subject, as target contrast increases the magnitudes of both facilitatory and suppressive effects diminish. The averaged data from all six subjects shows high levels of facilitation and suppression at low contrast, and gradual decline in the magnitude of both effects as target contrast increases. At low target contrast, the magnitudes of both effects, i.e., around 80% for facilitatory effects and around -60% for suppressive effects, are by far beyond previously reported magnitudes of contextual interactions as measured in visual sensitivity studies (Polat & Sagi, 1993; Kapadia et al., 1995). At higher contrasts, the target is always seen.

In the second experiment, I added a symmetric arrowhead at the other end of the target (Figure 38A), such as to resemble the full Müller-Lyer configuration. I also added a new target contrast level, 1.86%, and collected data from four subjects.

Figure 37 *Experiment 1.* (A) One-arrowhead stimulus schematic representation. (B) % increase in target detectability as a function of target contrast. Target detection is improved in the inward-pointing-arrowhead condition and it is impaired in the outward-pointing-arrowhead condition. Both suppressive and facilitatory effects diminish with increasing target contrast. The arrowhead angles were 54° (outward pointing arrowhead) and 286° (inward pointing arrowhead). I adjusted the proportion of positive responses in each of the four contrast conditions to compensate for guessing by presenting null conditions in which no target was presented (Kapadia et al., 1995). Performance levels at contrast 0.43% for subjects G. L. and L. T. are not presented because these subjects did not see the target at all at this contrast. (C) % increase in target detectability as a function of target contrast (average data). Percentages are calculated relative to the no-arrowhead condition.



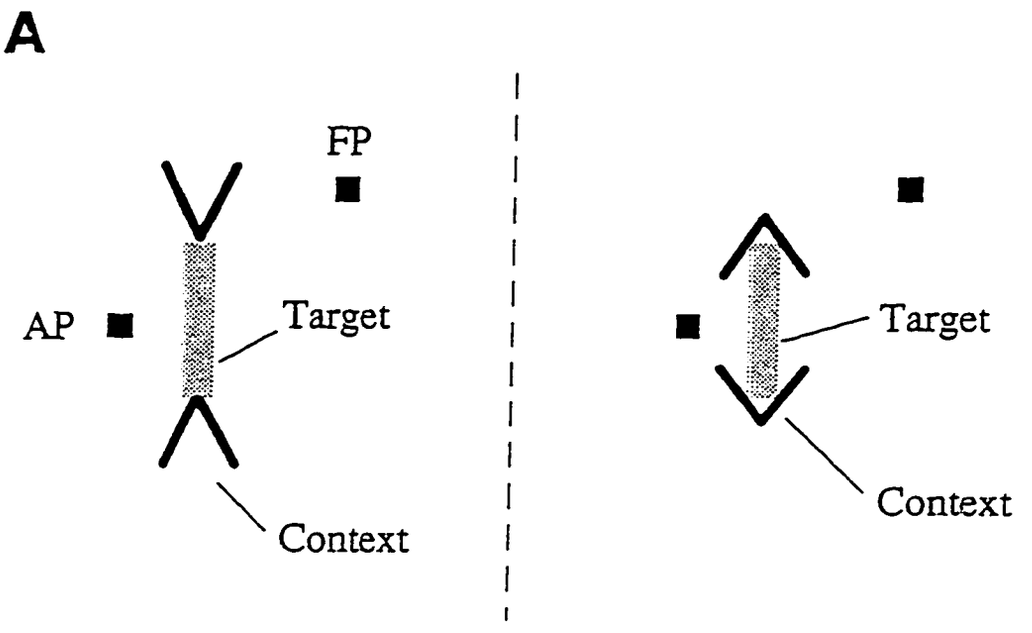


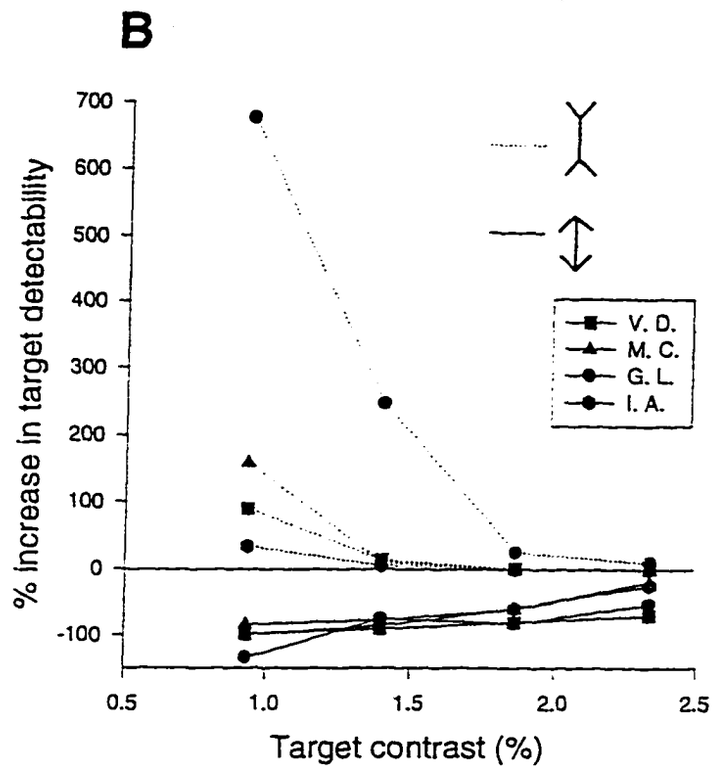


The data, summarized in Figures 38B and 38C, show that the facilitatory and suppressive effects found in the first experiment (see Figure 37) are amplified when the second arrowhead is added. It is noteworthy that these variations in target detectability match qualitatively the results obtained with the Müller-Lyer illusion, in which a high contrast bar appears long or short when it is flanked by inward- or outward-pointing arrowheads. Performance levels from the contrast 0.43% condition are not shown because three out of four subjects did not detect the target at all.

In the third experiment, I varied the half angle between the arrowhead fins (27°, 53°, 90°, 117°, and 143°) and the distance between the tip of the arrowhead and the target (30′, 1°, and 1°30′ of arc below the target base). The target contrast levels are identical to those used in Experiment 1. The stimulus configuration (angle and distance) is shown schematically in Figure 39A. Each of the four subjects was asked to detect the target bar when these 60 conditions (5 angles x 3 distances x 4 contrasts) were randomly intermixed. For short distances between the target and the arrowhead, the effect of acute half angles is suppressive, whereas the effect of obtuse half angles is facilitatory, at all target contrast levels. Furthermore, the degree of facilitation and suppression increases monotonically with the decrease in arrowhead angle, and diminishes as target contrast increases (Figure 39B). However, when the distance between the target and the context is increased, both inhibitory and excitatory effects diminish drastically in strength (Figure 39B). An unexpected result is that when distance is further increased (e.g., 1°30′ of arc below the target base) the influence of the outward-pointing arrowhead changes from suppression to facilitation (Figure 39B); the facilitatory effect of the outward-pointing arrowhead becomes stronger than the facilitatory effect of the inward-pointing arrowhead (which provided the maximum facilitation at short distances), see Figure 39C. This result matches qualitatively the Müller-Lyer extent illusion when the distance between the horizontal shaft and the two arrowheads increases (Yanagisawa, 1939; Fellows, 1967).

Figure 38 *Experiment 2.* (A) Two-arrowhead stimulus schematic representation. I added a second, symmetric, arrowhead at the other end of the target (the same stimulus conditions as in Experiment 1). (B) % increase in target detectability as a function of target contrast. Both suppressive and facilitatory effects are amplified compared to the one-arrowhead condition; the effects diminish with increasing target contrast. (C) % increase in target detectability as a function of target contrast (average data). In this analysis data from subject G. L., who showed levels of suppression and facilitation, were eliminated because they were out of the range in comparison with other subjects. If data from subject G. L. is included the overall effect is amplified. Percentages are calculated relative to the no-arrowhead condition.





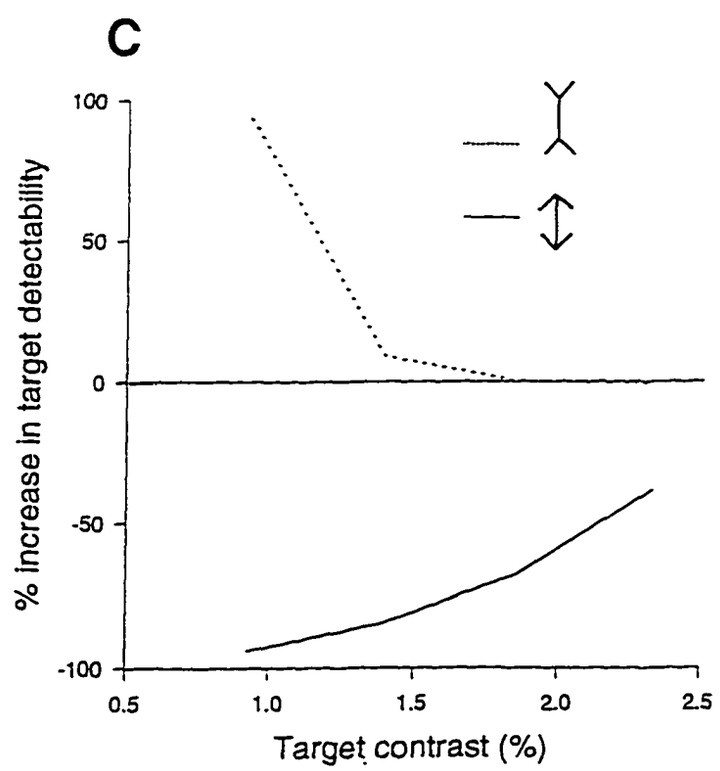
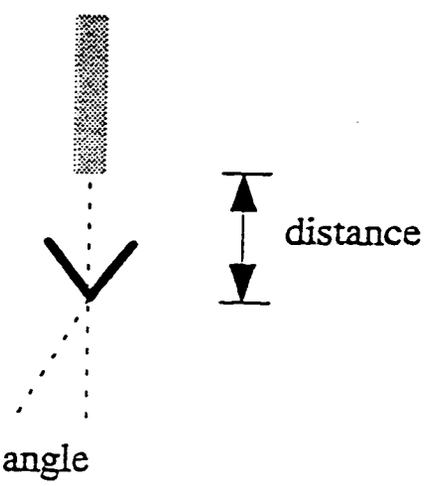
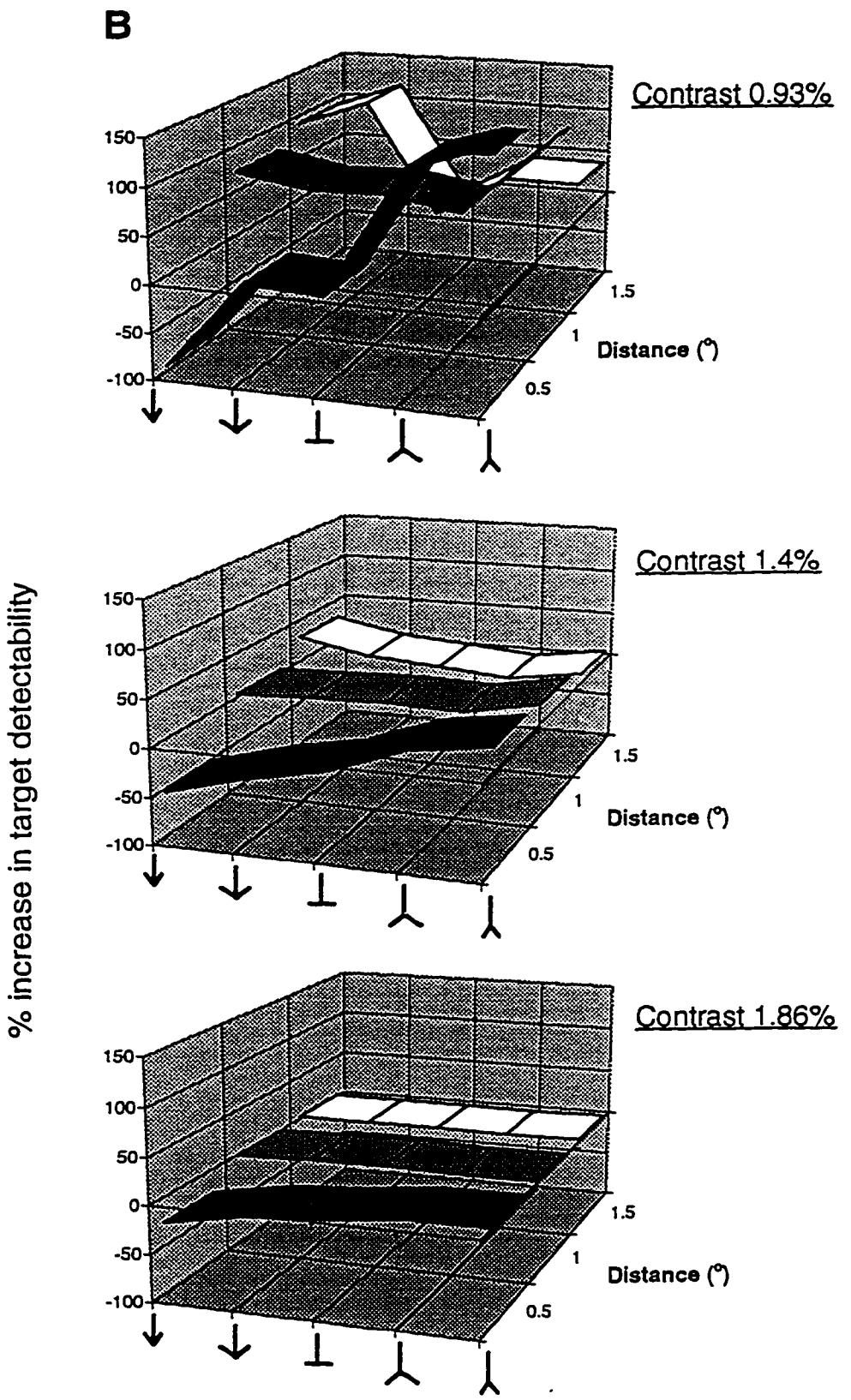
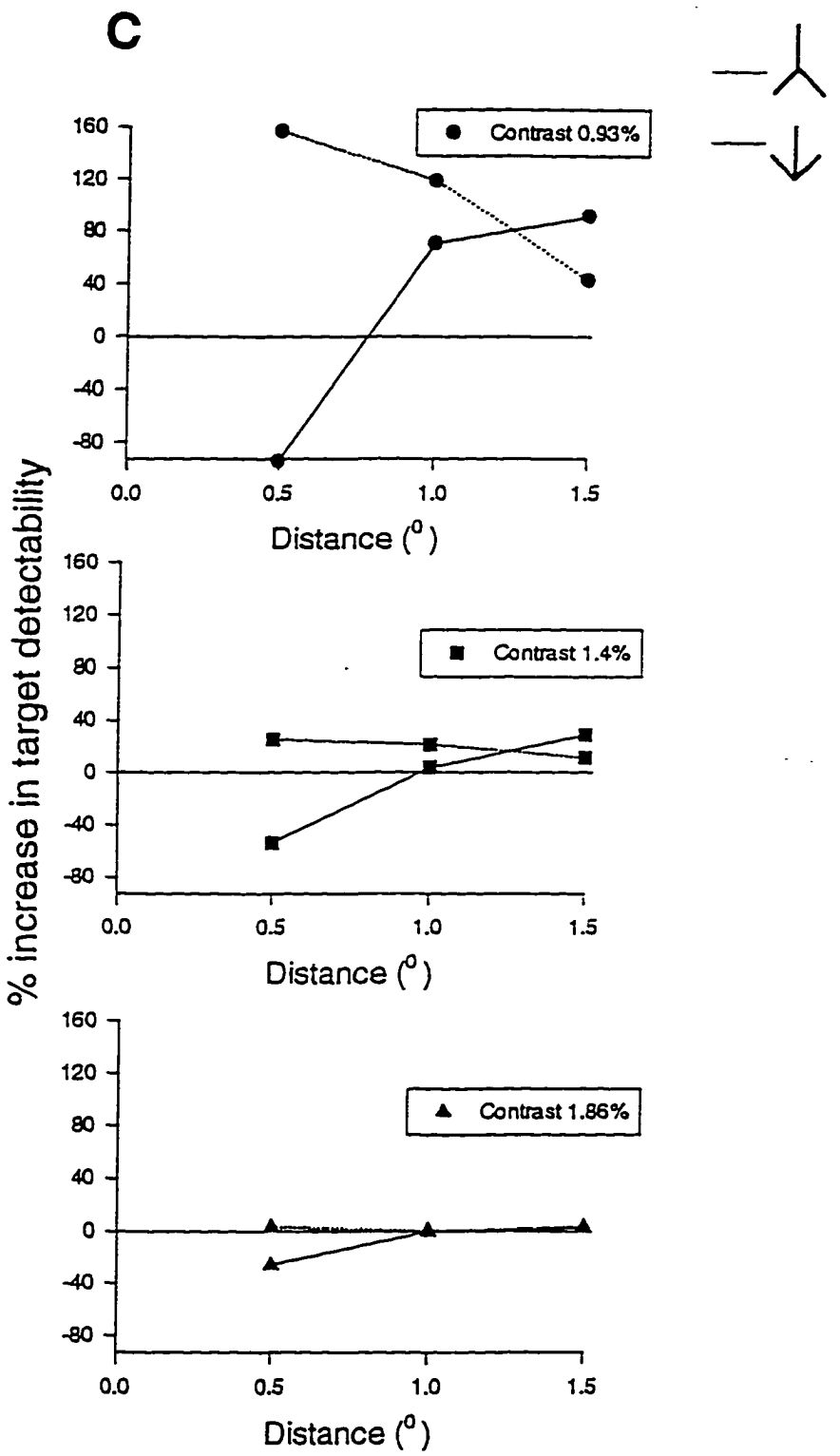


Figure 39 *Experiment 3.* (A) One-arrowhead stimulus schematic representation. I varied the half angle between the arrowhead fins (27°, 53°, 90°, 117°, and 143°) and the distance between the tip of the arrowhead and the target (30′, 1°, and 1°30′ of arc below the target base). The target contrast levels are identical to those used in Experiment 1. (B) % increase in target detectability as a function of arrowhead angle, arrowhead distance, and target contrast (average data). The suppressive effects diminish and the facilitatory effects increase as the arrowhead half angle becomes more obtuse. The effect of the outward-pointing-arrowhead changes from strong suppression to strong facilitation as the distance between target and context increases. Performance levels from the contrast 0.43% condition are not shown because three out of four subjects did not detect the target at all. Both suppressive and facilitatory effects diminish with increasing target contrast. Legends – black: 0.5°, gray: 1°, white: 1.5°. (C) % increase in target detectability as a function of arrowhead distance and target contrast (average data). The analysis was performed for arrowhead angles identical to those in Experiment 1. Both suppressive and facilitatory effects diminish with increasing distance between target and context. At distance 1°30′ the facilitatory effect of the outward-pointing-arrowhead becomes stronger than the facilitatory effect of the inward-pointing-arrowhead (the two curves intersect). Percentages are calculated relative to the no-arrowhead condition.

A







The model

The surround effects reported here favor 'low level' mechanisms of the Müller-Lyer illusion. In fact, I found a correlate of this illusion in the detectability domain. To explain these results, and related phenomena (Polat & Sagi, 1993; Kapadia et al., 1995), I have analyzed the properties of a population model of excitatory and inhibitory orientation detectors in visual cortex. This level of explanation was chosen because data from psychophysical studies of contrast detection threshold in human observers are correlated with single-cell recording in awake monkeys (Kapadia et al., 1995) and with visual evoked potentials in human observers (Polat & Sagi, 1993), thus suggesting a physiological basis for surround modulation. Figure 40A shows a schematic representation of the model. I model populations of local excitatory and inhibitory cells (cells which are activated by a stimulus presented in the classical receptive field, which is the region over which a stimulus evokes a suprathreshold response) and excitatory cells in the surround (also known as non-classical receptive field), and focus on local population dynamics. Sensory input indirectly activates both cortical excitatory and inhibitory cells through feedforward connections. The horizontal bar (center stimulus) is projected onto both local excitatory and inhibitory populations within the classical receptive field (CRF). The arrowhead context stimulus is projected onto excitatory cells in the surround (we do not model surround inhibitory cells because their intrinsic connections are purely local). Members of the local excitatory population are interconnected by recurrent excitatory synapses, and members of the inhibitory population are interconnected by recurrent inhibitory synapses. Local excitatory cells excite inhibitory cells, which in turn inhibit local excitatory cells.

Excitatory and inhibitory populations are modeled separately as threshold units with firing rates decribed by:

$$\frac{dE}{dt} = r_e(N_{fe}w_{fe}INP + N_{re}w_{re}E + N_{ee}w_{ee}S)(1 - E) - N_{li}w_{li}IE \qquad (17)$$

for the excitatory population, and

$$\frac{dI}{dt} = r_i(N_{fi}w_{fi}INP + N_{le}w_{le}E + N_{ei}w_{ei}S)(1 - I) - N_{ri}w_{ri}I \qquad (18)$$

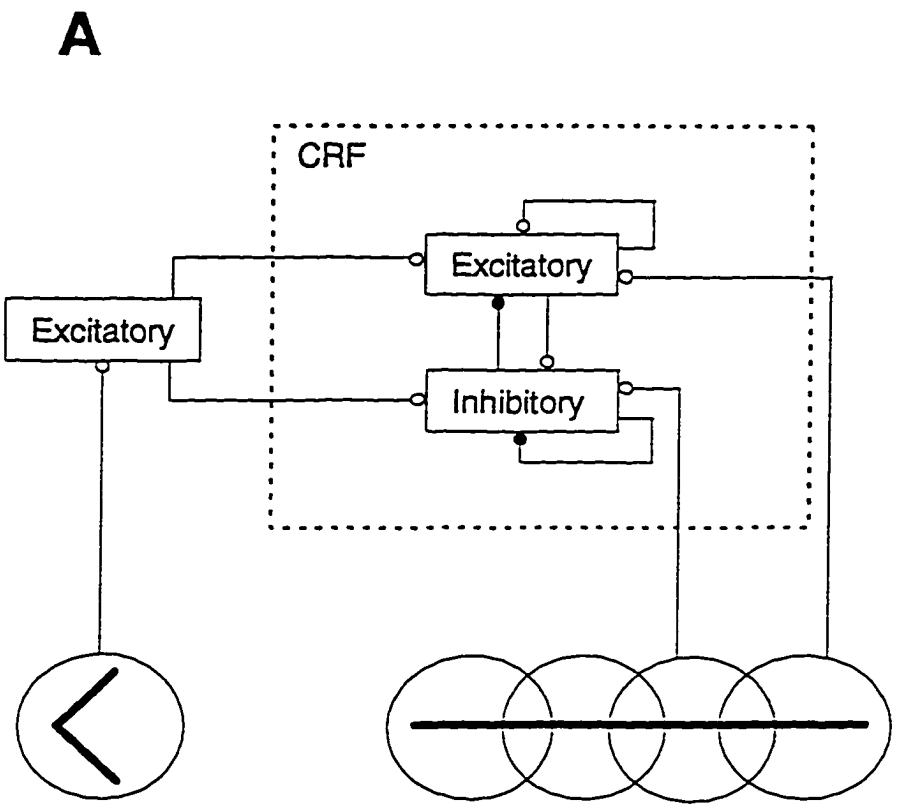
for the inhibitory population. Each unit becomes active when its firing rate is greater than a threshold (inhibitory cells have a higher threshold, 0.8, than excitatory cells, 0.3). Model parameters were chosen such that to ensure a ratio around 4:1 between excitatory and inhibitory cells. Inhibitory cells fire at a higher rate (rate constant $r_i = 0.1$) than excitatory cells (rate constant $r_e = 0.04$). The parameters are: $N_{fe} = N_{fi} = 50$ - number of feedforward projections to excitatory (fe) and inhibitory (fi) cells; $N_{re} = 50$ - number of projections from local (recurrent) excitatory cells; $N_{ri} = 10$ - number of projections from local (recurrent) inhibitory cells; $N_{li} = 10$ - number of projections from local inhibitory cells to local excitatory cells; $N_{le} = 40$ - number of projections from local excitatory cells to local inhibitory cells; $N_{ee} = 1000$ - number of long-range projections from surround cells to excitatory cells; $N_{ei} = 250$ - number of long-range projections from surround cells to inhibitory cells; $w_{fe} = w_{fi} = 0.08$ - connection strengths of feedforward projections to both excitatory (fe) and inhibitory (fi) cells; $w_{re} = w_{ri} = 0.005$ - connection strengths of recurrent excitatory (re) and inhibitory (ri) projections; $w_{le} = w_{li} = 0.1$ - connection strengths of local excitatory projections to inhibitory (le) cells and local inhibitory projections to excitatory (li) cells; $w_{ee} = w_{ei} = 0.1$ - connection strengths of long-range excitatory projections to both excitatory (ee) and inhibitory (ei) cells; INP is the total instantaneous input to both inhibitory and excitatory cells; $S = 0.02s \cdot exp^{-0.08(D^2+O^2)}$ - total input from excitatory cells in the surround, where D and O are the distance and orientation mismatch between center and surround, and s is the firing rate of the surround cells.

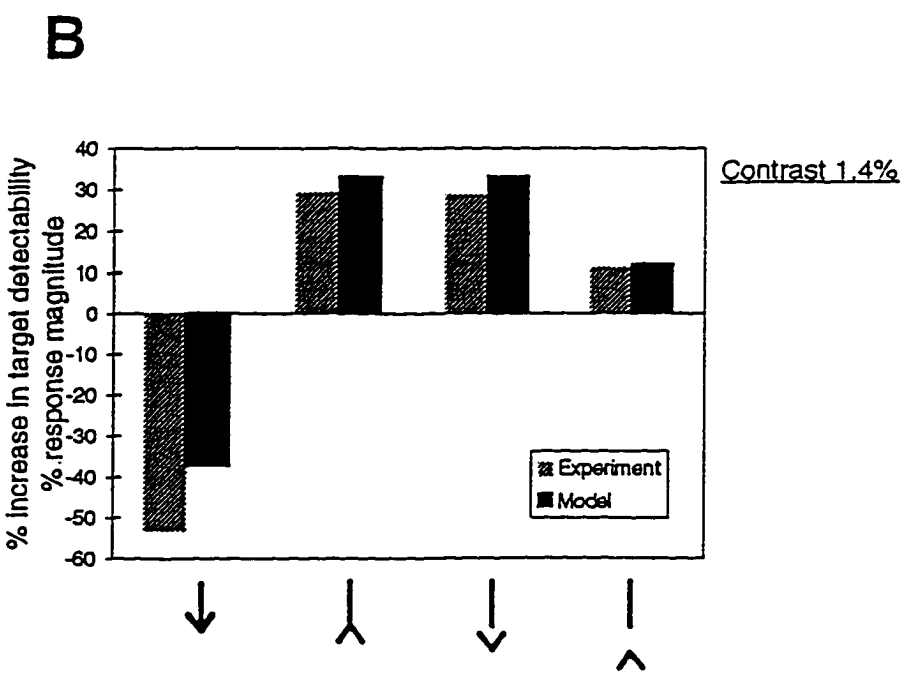
Surround effects are modeled using long-range excitatory connections on both local populations of excitatory and inhibitory cells. The strength of long-range horizontal connections is both orientation and distance-dependent, with the synaptic strengths decaying exponentially with the increase in the relative orientation preference between surround and center cells and with the increase in the distance

between center and surround. Surround population is modeled by a similar equation to those used to model local excitatory and inhibitory cells, except that cells in the surround only receive feedforward projections with a density of 50.

Figure 40B shows key experimental data in comparison with model predictions (see explanations in the figure legends). For short distances between center and surround (e.g., 30′ in our second experiment) the outward-pointing arrowhead excites strongly both local excitatory and inhibitory cells within CRF (we define the CRF as the region covering the target end which is closer to the arrowhead). Because the inhibitory cells typically fire at a higher rate and they have a higher threshold than excitatory cells the net effect is inhibitory (Weliky et al., 1995; Toth et al., 1996). and detectability of the target stimulus is impaired. In contrast, when the surround stimulus is the inward-pointing arrowhead, which is located outside CRF. the more distant excitatory cells in the surround activate only weakly the local inhibitory cells. The net effect is excitatory, and detectability of the target stimulus is enhanced. Both effects are orientation-dependent; because the strengths of long-range connections decay with the increase in the relative orientation between target and arrowhead there is both stronger suppression and facilitation when the arrowhead angle is more acute. However, when the distance between the horizontal bar and arrowhead increases the surround influence weakens, such that if distance is further increased the higher threshold local inhibition is shut off completely. In this situation, the influence of the arrowhead reverses; there is stronger facilitation by the outward-pointing arrowhead (surround is closer to the center, e.g., 1° or 1°30′ in our experiments) than by the inward-pointing arrowhead (in the simulations I calculate the distance between center and surround by estimating the distance between the center of each arrowhead fin and the closest end of the horizontal bar).

Figure 40 (A) Schematic representation of the population model. Excitatory connections between populations are shown as open circles: inhibitory ones, as filled circles. The horizontal bar stimulus is projected onto both local excitatory and inhibitory populations within the classical receptive field (CRF). The arrowhead context stimulus is projected onto excitatory cells in the surround. (B) Model predictions (%response magnitude) are compared with key experimental results (% increase in target detectability). Percentages are calculated relative to the no-arrowhead condition. Stimulus configurations are represented as angle/distance (from left to right): 27°/30′, 143°/30′, 27°/1°30′, 143°/1°30′; Target contrast is 1.4%.





Discussion

Three conclusions can be drawn from the present studies: (a) As a function of target contrast, surround orientation and spatial offset there is a continuum of surround modulations ranging from strong inhibition to strong excitation. Previous results (Polat & Sagi, 1993; Kapadia et al., 1995) reported both facilitatory and weak suppressive effects induced by a suprathreshold bar that flanks a collinear subthreshold target, and by Gabor patches surrounding a target Gabor patch of similar orientation and spatial frequency. (b) The present findings are consistent with short and long-range receptive field interactions in primary visual cortex (Weliky et al., 1995; Toth et al., 1996) and suggest a physiological basis for surround modulation. A population model of orientation detectors in visual cortex explains the present results as a byproduct of orientation and distance effects of long-range horizontal connections. The magnitudes of both facilitatory and suppressive effects that we found depend on the collinearity between target and surround. This is consistent with recent studies in primary visual cortex of tree shrews (Fitzpatrick, 1996) and squirrel monkeys (Blasdel et al., 1995) showing a link between orientation preference and the axial symmetry of lateral connections in the upper layers of striate cortex. Relevant to these investigations is the fact that in our pilot studies (Dragoi & Lockhead – unpublished data) we found that when the arrowhead is displaced laterally with respect to the symmetry axis of the target stimulus the effects on target detectability are weak, or none. (c) The effect of the arrowhead surround on the detection of a briefly exposed target stimulus (200 ms, which is less than a saccade reaction time) parallels the classical Müller-Lyer effect for a freely viewed pattern: facilitation in the case of the inward-pointing-arrowhead configuration (enhanced detectability in the present experiment and longer shaft in the Müller-Lyer effect) and suppression in the case of the outward-pointing-arrowhead configuration (diminished detectability in our experiment and shorter shaft in the Müller-Lyer effect). These spatial interactions for target detectability and for the Müller-Lyer illusion are nonlinear functions of contrast, orientation,

and distance. I disambiguate the concept of 'confusion' typically associated with the perceived extent in Müller-Lyer illusion suggesting that the same 'low level' mechanism responsible for large modulations in visual sensitivity at low contrasts is involved in 'higher level' processing when perceiving exact location of horizontal lines flanked by two arrowheads in the Müller-Lyer configuration.

Appendix

All the units described in the text vary between 0 and 1. All the computer simulations have been performed equating the integration step (1.2 for all the simulated paradigms) with a formal time unit (1 second). A reinforcement of variable magnitude is applied according to all the simulated reinforcement schedules. The model's predictions are consistent with experimental data for a large range of coefficients. Parameter values used in all simulations are $\alpha_1 = 1 \cdot 10^{-4}$, $\alpha_2 = 0.05$, $\alpha_3 = 0.2$, $\alpha_4 = \alpha_6 = \alpha_9 = \alpha_{11} = 0.1$, $\alpha_5 = 0.5$, $\alpha_7 = 8 \cdot 10^{-3}$, $\alpha_8 = \alpha_{10} = 5 \cdot 10^{-3}$, $\alpha_{12} = 5 \cdot 10^{-5}$, $\alpha_{13} = 1.5 \cdot 10^{-4}$, $\alpha_{14} = 0.03$.

References

[1] Ahmed, B. A., Anderson, J. C., Douglas, R. J., Martin, K. A. C., & Nelson, J. C. (1994). Polyneuronal innervation of spiny stellate neurons in cat visual cortex. *Journal of Comparative Neurology*, **341**, 39-49.

[2] Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, **82**, 463-496.

[3] Ainslie, G., & Herrnstein, R. J. (1981). Preference reversal and delayed reinforcement. *Animal Learning & Behavior*, **9**, 476-482.

[4] Amsel A. (1962). Frustrative nonreward in partial reinforcement and discrimination learning. *Psychological Review*, **69**, 306-328.

[5] Bacon, W. E. (1962). Partial-reinforcement extinction effect following different amounts of training. *Journal of Comparative and Physiological Psychology*, **55**, 998-1003.

[6] Belke, T. W. (1992). Stimulus preference and the transitivity of preference. *Animal Learning & Behavior*, **20**, 401-406.

[7] Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, **32**, 269-281.

[8] Benefield, R., Oscos, A., & Ehrefreund, D. (1974). Role of frustration in successive positive contrast effect. *Journal of Comparative and Physiological Psychology*, **86**, 648-651.

[9] Bernheim, J. W., & Williams, D. R. (1967). Time-dependent contrast effects in a multiple schedule of food reinforcement. *Journal of the Experimental Analysis of Behavior*, **10**, 243-249.

[10] Black, R. W. (1968). Shifts in magnitude of reward and contrast effects in instrumental conditioning. *Psychological Review*, **75**, 114-126.

[11] Blasdel, G., Obermayer, K. & Kiorpes, L. (1995). Organization of ocular dominance and orientation columns in the striate cortex of neonatal macaque monkeys. *Visual Neurosci.*, **12**, 589-603.

[12] Blough, D. S. (1983). Alternative accounts of dimensional stimulus control. In M. Commons, R. Herrnstein, & A. Wagner (Eds.), *Quantitative analyses of behavior* (Vol. 4, pp. 59-72). Cambridge, MA: Ballinger.

[13] Bower, G. H. (1961). A contrast effect in differential conditioning. *Journal of Experimental Psychology*, **62**, 196-199.

[14] Burstein, K. R. (1967). Spontaneous Recovery: A (Hullian) noninhibition interpretation. *Psychonomic Science*, **7**, 389-390.

[15] Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning.* New York: Wiley.

[16] Capaldi, E. J. (1958). The effect of different amounts of training on the resistance to extinction of different patterns of partially reinforced responses. *Journal of Comparative Physiological Psychology*, **51**, 367-371.

[17] Capaldi, E. J. (1967). Sequential versus nonsequential variables in partial delay of reward. *Journal of Experimental Psychology*, **74**, 161-166.

[18] Capaldi, E. J. (1971). Memory and learning: a sequential viewpoint. In W. K. Honig and P. H. R. James (Eds.), *Animal Memory* (pp. 111-154). New York: Academic Press.

[19] Capaldi, E. J., & Kassover, K. (1970). Sequence, number of nonrewards, anticipation, and intertrial interval in extinction. *Journal of Experimental Psychology*, **84**, 470-476.

[20] Capaldi, E. J., & Lynch, D. (1967). Repeated shifts in reward magnitude: Evidence in favor of an associational and absolute (noncontextual) interpretation. *Journal of Experimental Psychology*, **75**, 226-235.

[21] Capaldi, E. J., & Minkoff, R. (1967). Reward schedule effects at a relatively long intertrial interval. *Psychonomic Science*, **9**, 169-170.

[22] Capaldi, E. J., & Stevenson, H. W. (1957). Response reversal following different amounts of training. *Journal of Comparative Physiological Psychology*. **50**, 195-198.

[23] Chiang, C. (1968). A new theory to explain geometrical illusions produced by crossing lines. *Perception & Psychophysics*, **3**, 174-176.

[24] Chung, S. H., & Herrnstein, R. J. (1967). Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, **10**, 67-74.

[25] Churchland, P. S., Ramachandran, V. S., Sejnowski, T. J. (1993). A critique of pure vision. In Large-scale Neuronal Theories of the Brain, C. Koch and J. L. Davis Eds. (MIT Press, Cambridge MA), pp 23-60.

[26] Cohen, S. L., Riley, D. S., & Weigle. P. A. (1993). Tests of behavior momentum in simple and multiple schedules with rats and pigeons. *Journal of the Experimental Analysis of Behavior*, **60**, 255-291.

[27] Cox, W. M. (1975). A review of recent incentive contrast studies involving-discrete trial procedures. *The Psychological Record*, **25**, 373-393.

[28] Crespi, L. P. (1952). Quantitative variation of incentive and performance in the white rat. *American Journal of Psychology*, **55**, 467-517.

[29] Crum, J., Brown, W. L., & Bitterman, M. E. (1951). The effect of partial and delayed reinforcement on resistance to extinction. *American Journal of Psychology*, **64**, 228-237.

[30] Daly, H. B., & Daly, J. T. (1982). A mathematical model of reward and aversive nonreward: its application in over 30 appetitive learning situations. *Journal of Experimental Psychology: General*, **111**, 441-480.

[31] Davenport, J. W. (1962). The interaction of magnitude and delay of reinforcement in spatial discrimination. *Journal of Comparative Physiological Psychology*, **55**, 267-273.

[32] Davis, D. G. S, & Staddon, J. E. R. (1990). Memory for reward in probabilistic choice: Markovian and non-Markovian properties. *Behaviour*, **114**, 37-64.

[33] Davis, D. G. S., Staddon, J. E. R., Machado, A., & Palmer, R. (1993). The process of recurrent choice. *Psychological Review*, **100**, 320-341.

[34] Davison, M., & McCarthy, D. (1988). *The matching law.* Hillsdale, N. J.: Erlbaum.

[35] Di Lollo, V., & Beez, V. (1966). Negative contrast effect as a function of magnitude of reward decrement. *Psychonomic Science*, **5**, 99-100.

[36] Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. C., & Suarez, H. H. (1995). Recurrent excitation in neocortical circuits. *Science*, **269**, 981-985.

[37] Dragoi, V., & Staddon, J. E. R. (1993). A competitive neural network model for the process of recurrent choice. In M. C. Mozer, P. Smolensky, D. S. Touretzky, J. L. Elman, & A. S. Weigend (Eds.), *Proceedings of the 1993 Connectionist Models Summer School* (pp. 65-73). Hillsdale, NJ: Erlbaum Associates.

[38] Dragoi, V. (1995). Cortical dynamics of form perception: Geometrical illusions and aftereffects. Proceedings of the World Congress on Neural Networks, **3**, 275-278 (1995).

[39] Dragoi, V., & Wolbarsht, M. L. (1995). Context-dependent form perception in visual cortex. Investigative Ophthalmology and Visual Science, **36**, 473 (1995).

[40] Dragoi, V. (1997). A dynamic theory of acquisition and extinction in operant learning. *Neural Networks*, **10**, 201-229.

[41] Dragoi, V. (in press). A model of contextual interactions in primary visual cortex: Examining the influence of corticogeniculate feedback. In: Computational Neuroscience '96. (Bower JM, ed.). Plenum Press.

[42] Dragoi, V., & Staddon, J. E. R. (in review). The dynamics of operant conditioning. *Psychological Review.*

[43] Dreyfus, L. R. (1991). Local shifts in relative reinforcement rate and time allocation on concurrent schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 486-502.

[44] Erlebacher, A. & Sekuler, R. (1969). Explanation of the Müller-Lyer illusion: Confusion theory examined. *Journal of Experimental Psychology*, **80**, 462-467.

[45] Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, **62**, 145-154.

[46] Ettinger, R. H., & Staddon, J. E. R. (1982). Behavioral competition, component duration and multiple-schedule contrast. *Behaviour Analysis Letter*, **2**, 31-38.

[47] Fellows, B. J. (1967). Reversal of the Müller-Lyer illusion with changes in the length of the inter-fins line. *Quarterly Journal of Experimental Psychology*, **19**, 208-214.

[48] Fisher, G. H. (1967). A common principle relating to the Müller-Lyer and Ponzo illusions. *American Journal of Psychology*, **80**, 626-631.

[49] Fitzpatrick, D. (1996). The functional organization of local circuits in visual cortex: insights from the study of tree-shrew striate cortex. *Cerebral Cortex*,

6, 329-341.

[50] Franchina, J. J., & Brown, T. S. (1971). Reward magnitude shifts effects in rats with hippocampal lesions. *Journal of Comparative Physiological Psychology*, **76**, 365-370.

[51] Fries, W., Albus, K., & Creutzfeldt, O. D. (1977). Effects of interacting visual patterns on single cell responses in cat's striate cortex. *Vision Research*, **17**, 1001-1008.

[52] Gallistel, C. R. (1990). *The organization of learning*. MIT Press.

[53] Garey, L. J., & Powell, T. P. S. (1971). An experimental study of the termination of the lateral geniculo-cortical pathway in the cat and monkey. *Proc. Royal Soc. London [Biol]*, **179**, 41-63.

[54] Gibbon, J. (1995). Dynamics of time matching: Arousal makes better seem worse. *Psychonomic Bulletin & Review*, **2**, 208-215.

[55] Gibbon, J., Church, R. M., Fairhurst, S., & Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. *Psychological Review*, **95**, 102-114.

[56] Gilbert, C. D., & Wiesel, T. N. (1979). Morphology and intracortical projections of functionally identified neurons in cat visual cortex. *Nature*, **280**, 120-125.

[57] Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal connections and corticocortical connections in cat visual cortex. *Journal of Neuroscience*, **9**, 2432-2442.

[58] Gilbert, C. D., & Wiesel, T. N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research*, **30**, 1689-1701.

[59] Gillam, B. (1990). In *The Perceptual World*, I. Rock Ed. (W. H. Freeman and Co), 164-176.

[60] Gray, J. A. (1971). *The psychology of fear and stress.* Weidenfeld and Nicholson.

[61] Gray, J. A. (1982). *The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system.* Oxford University Press, New York.

[62] Greene, E. & Nelson, B. (1997). Evaluating Müller-Lyer effects using single fin-set configurations. *Perception & Psychophysics*, **59**, 293-312.

[63] Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, **199**, 678-680.

[64] Grieve, K. L., & Sillito, A. M. (1995). Differential properties of cells in the feline primary visual cortex providing the corticofugal feedback to the lateral geniculate nucleus and visual claustrum. *Journal of Neuroscience*, **15**, 4868-4874.

[65] Grinvald, A., Lieke, E. E., Frostig, R. D., & Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque primary visual cortex. *Journal of Neuroscience*, **14**, 2545-2568.

[66] Grossberg, S. (1972). A neural theory of punishment and avoidance, II: Quantitative theory. *Mathematical Biosciences*, **15**, 39-67.

[67] Grossberg, S. (1981). Psychophysiological substrates of schedule interactions and behavioral contrast. *SIAM-AMS Proceedings*, **13**, 157-186.

[68] Grossberg, S. (1982). Processing of expected and unexpected events during conditioning and attention: a psychophysiological theory. *Psychological Review*, **89**, 529-572.

[69] Gutman, A. (1977). Positive contrast, negative induction, and inhibitory stimulus control in rat. *Journal of the Experimental Analysis of Behavior*, **27**, 219-233.

[70] Halliday, M. S., & Boakes, R. A. (1971). Behavioral contrast and response independent reinforcement. *Journal of the Experimental Analysis of Behavior*, **16**, 429-434.

[71] Harvey, A. R. (1980). A physiological analysis of subcortical and commissural projections of area 17 and 18 of the cat. *Journal of Physiology (London)*, **302**. 507-534.

[72] Hata, Y., Tsumoto, T., Sato, H., Hagihara, K., & Tamura, H. (1988). Inhibition contributes to orientation selectivity in visual cortex of cat. *Nature*, **335**, 815-817.

[73] Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of Experimental Analysis and Behavior*, **4**, 267-272.

[74] Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, **13**, 243-266.

[75] Herrnstein, R. J., & Vaughan, W. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to Action: The Allocation of Individual Behavior* (pp. 143-146). San Diego, CA: Academic Press.

[76] Hinson, J. M., & Staddon, J. E. R. (1983). Matching, maximizing and hill climbing. *Journal of the Experimental Analysis of Behavior*, **40**, 321-31.

[77] Hirsch, J. A., & Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *Journal of Neuroscience*, **11**, 1800-1809.

[78] Hornung, J. P., & Garey, L. J. (1981). The thalamic projection to the cat visual cortex: ultrastructure of neurons identified by Golgi impregnation or

retrograde horseradish peroxidase transport. *Neurosience*, **6**, 1053-1068.

[79] Hooper, R. (1967). Variables controlling the overlearning reversal effect (ORE). *Journal of Experimental Psychology*, **73**, 612-619.

[80] Horner, J. M, & Stadon, J. E. R. (1987). Probabilistic choice: A simple invariance. *Behavioural Processes*, **15**, 59-92.

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

[82] Kacelnik, A., & Krebs, J. R., & Ens, B. (1987). Foraging in a changing environment: an experiment with starlings (sturnus vulgaris). In M. L. Commons, A. Kacelnik & S. J. Shettleworth (Eds.), *Quantitative Analyses of Behavior VI: Foraging* (pp. 63-87). Hillsdale, NJ: Erlbaum.

[83] Kapadia, M. K., Ito, M., Gilbert, C. D.& Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron*, **15**, 843-856.

[84] Kelly, J. S., & Renaud, L. P. (1974). Physiological identification of inhibitory interneurons in the feline pericruciate cortex. *Neuropharmacology*, **13**, 463-474.

[85] Killeen, P. (1968). On the measurement of reinforcement frequency in the study of preference. *Journal of the Experimental Analysis of Behavior*, **11**, 263-269.

[86] Killeen, P. (1970). Preference for fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, **14**, 127-131.

[87] Killeen, P. (1994). Mathematical principles of reinforcement. *Behavioral and Brain Sciences*, **17**, 105-172.

[88] Kisvarday, Z. F., Martin, K. A. C., Freund, T. F., Magloczky, Z., Whiteridge, D., & Somogyi, P. (1986). Synaptic target of HRP-filled layer III pyramidal cells in the cat striate cortex. *Experimental Brain Research*, **64**, 541-552.

[89] Klopf, A. H. (1988). A neuronal model of classical conditioning. *Psychobiology*, **16**, 85-125.

[90] Knierim, J., & Van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of alert macaque monkey. *Journal of Neurophysiology*, **67**, 961-980.

[91] Lea, S. E. G., & Dow, S. M. (1984). The integration of reinforcements overtime. *Annals of the New York Academy of Sciences*, **423**, 269-277.

[92] LeVay, S. M. (1986). Synaptic organization of claustral and geniculate afferents to the visual cortex of the cat. *Journal of Neuroscience*, **6**, 3564-3575.

[93] Livingston, M. S., & Hubel, D. H. (1984). Specificity of intrinsic connections in primate primary visual cortex. *Journal of Neuroscience*, **4**, 2830-2835.

[94] Luce, R. D. (1959). *Individual choice behavior: a theoretical analysis*. New York: Wiley.

[95] Mackintosh, N. J. (1969). Further analysis of the overtraining reversal effect. *Journal of Comparative and Physiological Psychology. Monograph*, **67**, 2, Part 2.

[96] Mackintosh, N. J. (1974). *The psychology of animal learning*. Academic Press.

[97] Malach, R., Amir, Y., & Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and *vivo* tangential biocytin injections in primate striate cortex. *Proc. Natl. Acad. Sci. USA*, **90**, 10469-10473.

[98] Mark, T. A., & Gallistel, C. R. (1994). Kinetics of matching. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 79-95.

[99] Mark, T. A., & Gallistel, C. R. (1995). The microstructure of matching. Proceedings of Psychonomics, oral presentation.

[100] Martin, K. A. C., & Whitteridge, D. (1984). Form, function, and intracortical projections of spiny neurons in the striate visual cortex of the cat. *Journal of Physiology*, **353**, 463-504.

[101] Maxwell, F. R., Calef, R. S., Murray, D. W., Shepard, D. C., & Norville, R. A. (1976). Positive and negative successive contrast effects following multiple shifts in reward magnitude under high drive and immediate reinforcement. *Animal Learning and Behavior*, **4**, 480-484.

[102] Mazur, J. E. (1992). Choice behavior in transition: development of preference with ratio and interval schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 364-378.

[103] Mazur, J. E. (1995). Development of preference and spontaneous recovery in choice behavior with concurrent variable-interval schedules. *Animal Learning and Behavior*, **23**, 93-103.

[104] Mazur, J. E., & Ratti, T. A. (1991). Choice behavior in transition: Development of preference in a free-operant procedure. *Animal Learning & Behavior*, **19**, 241-248.

[105] McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology, General*, **114**, 159-188.

[106] McClurkin, J. W., Optican, L. M., & Richmond, B. J. (1994). Cortical feedback increases visual information transmitted by monkey parvocellular lateral geniculate nucleus neurons. *Visual Neuroscience*, **11**, 601-617.

[107] McCormick, D. A., Connors, B. W., Lighthall, J. W., & Prince, D. A. (1985). Comparative electrophysiology of physiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *Journal of Neurophysiology*, **54**, 782-806.

[108] McEwen, D. (1972). The effects of terminal-link fixed-interval and variable-interval schedules on responding under concurrent chained schedules. *Journal of the Experimental Analysis of Behavior*, **18**, 253-261.

[109] McGuire, B., Gilbert, C. D., Rivlin, P. K., & Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *Journal of Comparative Neurology*, **305**, 370-392.

[110] Meyer, D. R. (1951). The effects of differential rewards on discrimination reversal learning by monkeys. *Journal of Experimental Psychology*, **41**, 268-274.

[111] Michalski, A., Gerstein, G. I., Czarkowska, J., & Tarnecki, R. (1983). Interactions between cat striate cortex neurons. *Experimental Brain Research*, **51**, 97-107.

[112] Mukherjee, P., Ozaki, T., & Kaplan, E. (1995). What controls the transfer of information through the LGN? *Soc. Neurosci. Abstr.*, **21**, 657.

[113] Mikulka, P. J., Lehr, R., & Pawlik, W. B. (1967). Effect of reinforcement schedules on reward shifts. *Journal of Experimental Psychology*, **74**, 57-61.

[114] Müller-Lyer, F. C. (1889). Optische Urteilstäuschungen [Optical illusions]. *Archiv für Anatomie und Physiologie: Physiologische Abteilung*, **2** (Suppl.), 263-270.

[115] Myerson, J., & Miezin, F. M. (1980). The Kinetics of choice: An operant systems analysis. *Psychological Review*, **87**, 160-174.

[116] Nelson, J. I. & Frost, B. J. (1981). Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Research*, **139**, 359-365.

[117] Nevin, J. A., & Shettleworth, S. J. (1966). An analysis of contrast effects in multiple schedules. *Journal of the Experimental Analysis of Behavior*, 9, 305-315.

[118] Nevin, J. A. (1988). Behavioral momentum and the partial reinforcement effect. *Psychological Bulletin*, **103**, 44-56.

[119] Pavlov, I. P. (1927). *Conditioned reflexes*. Oxford: Oxford University Press.

[120] Polat, U. & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research*, **33**, 993-999.

[121] Pressey, A. W. (1970). The assimilation theory applied to a modification of the Müller-Lyer illusion. *Perception & Psychophysics*, 8, 411-412.

[122] Reid, L. S. (1953). The development of noncontinuity behavior through continuity learning. *Journal of Experimental Psychology*, **46**, 107-112.

[123] Rensink, R. A. & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, **102**, 101-130.

[124] Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory*. New York: Appleton-Century-Crofts.

[125] Reynolds, B. (1961). Behavioral contrast. *Journal of Experimental Analysis of Behavior*, 4, 57-71.

[126] Richardson, R., & Campbell, B. A. (1992). Latent habituation of the orienting response in the preweanling rat. *Animal Learning & Behavior*, **20**, 416-426.

[127] Robbins, S. J. (1990). Mechanisms underlying spontaneous recovery in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 235-249.

[128] Robinson, J. O. (1972). *The Psychology of Visual Illusions.* Hutchinson University Library, London.

[129] Rockland, K. S., & Lund, J. S. (1982). Widespread periodic intrinsic connections in the tree shrew visual cortex. *Brain Research*, **169**, 19-40.

[130] Rozin, P. (1965). Temperature independence of an arbitrary temporal discrimination in the goldfish. *Science*, **149**, 561-563.

[131] Rudy, J. W. (1971). Sequential variables as determiners of the rat's discrimination of reinforcement events: effects of extinction performance. *Journal of Comparative Physiological Psychology*, **77**, 476-481.

[132] Russel, W. R. (1971). *The traumatic amnesia.* London: Oxford University Press.

[133] Schmajuk, N. A., & DiCarlo, J. J. (1992). Stimulus configuration, classical conditioning, and hippocampal function. *Psychological Review*, **99**, 268-305.

[134] Schmajuk, N. A. (in press). *Animal learning and cognition: A neural network approach.* Cambridge University Press.

[135] Schwartz, B., & Gamzu, E. (1977). Pavlovian control of operant behavior: an analysis of autoshaping and its implication for operant conditioning. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of Operant Behavior.* Englewood Cliffs, N.J.: Prentice Hall.

[136] Sgro, J. A., & Weinstock, S. (1963). Effects of delay on subsequent running under immediate reinforcement. *Journal of Experimental Psychology*, **66**, 260-263.

[137] Shimp, C. P., & Wheatley, K. L. (1971). Matching to relative reinforcement frequency in multiple schedules with a short component duration. *Journal of the Experimental Analysis of Behavior*, **15**, 205-210.

[138] Sillito, A. M., Cudeiro, J., & Murphy, P. C. (1993). Orientation sensitive elements in the corticofugal influence on centre-surround interactions in the dorsal lateral geniculate nucleus. *Experimental Brain Research*, **93**, 6-16.

[139] Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis J. (1995). Visual cortical mechanisms deetcting focal orientation discontinuities. *Nature*. **378**, 492-496.

[140] Sillito, A. M., Jones, H. E., Gerstein, G. L., & West, D. C. (1995). Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature*, **369**, 479-482.

[141] Somers, D. C., Todorov, E. V., Siapas, A. G., & Sur, M. (1995). Vector-based integration of local and long-range information in visual cortex. *A. I. Memo No*, **1556** (Massachusetts Institute of Technology, Cambridge, MA).

[142] Spear, N. E. (1973). Retrieval of memory in animals. *Psychological Review*, **80**, 163-194.

[143] Spence, K. W. (1956). *Behavior theory and conditioning.* New Haven: Yale University Press.

[144] Squire, L. R., Slater, P. C., & Chace, P. M. (1975). Retrograde amnesia: Temporal gradient in very long term following electroconvulsive therapy. *Science*, **187**, 77-79.

[145] Staddon, J. E. R. (1983). *Adaptive behavior and learning.* Cambridge University Press.

[146] Staddon, J. E. R. (1988). Quasi-Dynamic choice models: Melioration and ratio invariance. *Journal of the Experimental Analysis and Behavior*, **49**, 2.

303-320.

[147] Staddon, J. E. R., & Higa, J. (in press). Multiple time scales in rate-sensitive habituation. *Psychological Review.*

[148] Staddon, J. E. R., & Hinson, J. M. (1978). Behavioral competition: a mechanism for schedule interactions. *Science*, **202**, 432-434.

[149] Staddon, J. E. R., & Simmelhag, V. G. (1971). The "superstition " experiment: A Reexamination of its implications for the principles of adaptive behavior. *Psychological Review*, **78**, 3-43.

[150] Staddon, J. E. R., & Zhang, Y. (1991). On the assignment-of-credit problem in operant learning. In M. L. Commons, S. Grossberg, & J. E. R. Staddon (Eds.), *Neural Network Models of Conditioning and Action* (pp. 279-293). Hillsdale, NJ: Erlbaum.

[151] Stemmler, M., Usher, M., & Niebur, E. (1995). Lateral interactions in primary visual cortex: A model bridging physiology and psychophysics. *Science.* **269**, 1877-1880.

[152] Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review*, **88**, 135-170.

[153] Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. W. Moore, (Eds.), *Learning and Computational Neuroscience: Foundations of Adaptive Networks.* Cambridge, MA: MIT Press.

[154] Tolman, E. C. (1932). *Purposive behavior in animals and men.* New York: Appleton-Century-Crofts.

[155] Tolman, E. C., & Honzik, C. H. (1930). Introduction and removal of reward, and maze performance in rats. *University of California Publications in Psychology*, **4**, 257-275.

[156] Toth, L. J., Rao, S. C., Kim, D. S., Somers, D. & Sur, M. (1996). Subthreshold facilitation and suppression in primary visual cortex revealed by intrinsic signal imaging. *Proceeding of the National Academy of Sciences USA*, **93**, 9869-9874.

[157] Toyama, K., Kimura, M., & Tanaka, K. (1981). Organization of cat visual cortex as investigated by cross-correlation technique. *Journal of Neurophysiology*, **46**, 202-214.

[158] Tyler, D. W., Wortz, E. C., & Bitterman, M. E. (1953). The effect of random and alternating partial reinforcement on resistance to extinction in the rat. *American Journal of Psychology*, **66**, 57-65.

[159] Vaughan, W. (1982). Choice and the Rescorla-Wagner model. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior, Vol 2: Matching and maximizing accounts.* Cambridge, Massachusetts: Ballinger.

[160] Vogel, J. R., Mikulka, P. J., & Spear, N. E. (1966). Effect of interpolated extinction and level of training on the "depression" effect. *Journal of Experimental Psychology*, **72**, 51-60.

[161] Weinstock, S. (1958). Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. *Journal of Experimental Psychology*, **46**, 151-158.

[162] Weisman, R. G. (1969). Some determinants of inhibitory stimulus control. *Journal of the Experimental Analysis of Behavior*, **12**, 443-450.

[163] Weliky, M., & Katz, L. C. (1994). Functional mapping of horizontal connections in developing ferret visual cortex: experiments and modeling. *Journal of Neuroscience*, **14**, 7291-7305.

[164] Weliky, M., Kandler, K., Fitzpatrick, D. & Katz, L. C. (1995). Pattern of excitation and inhibition evoked by horizontal connections in visual cortex

share a common relationship to orientation columns. *Neuron*, **15**. 541-552.

[165] Williams, B. A. (1983). Another look at contrast in multiple schedules. *Journal of the Experimental Analysis of Behavior*, **39**, 345-384.

[166] Williams, B. A., & Royalty, P. (1989). A test of the melioration theory of matching. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 99-113.

[167] Winfield, D. A., & Powell, T. P. S. (1983). Laminar cell counts and geniculocortical boutons in area 17 of cat and monkey. *Brain Research*, **227**, 223-229.

[168] Yanagisawa, N. (1939). Reversed illusion in the Müller-Lyer illusion. *Japanese Journal of Pychology*, **14**, 321-326.

Curriculum Vitae

Valentin Dragoi

Department of Psychology: Experimental
Duke University, Box 90086
Durham, NC 27708
tel: (919)-660-5677
fax: (919)-660-5726
email: valentin@psych.duke.edu

Education

* Duke University, Dept. of Psychology: Experimental (expected Ph.D. degree, September 1997)

* Polytechnic Institute of Iasi (Romania), Dept. of Computer Science, M.S., 1989

Professional Appointments

* Graduate Student, Duke University, Dept. of Psychology: Experimental, 1992-1997.

* Researcher, Institute for Information Sciences, Romanian Academy, Iasi Branch, 1990-1992.

* Instructor, Polytechnic Institute of Iasi (Romania), Dept. of Computer Science, 1990-1992.

* Computer Engineer, "IIRUC" Bucuresti (Romania), Iasi Branch, Dept. of Distributed Information, 1989-1990.

Awards and Distinctions

* *ARVO Retina Research Foundation* Travel Fellowship, 1997.

* *Marine Biological Laboratory* Fellowship, Methods in Computational Neuroscience Summer School, Woods Hole, MA, 1996.

* *National Academy of Sciences*, Sigma Xi, The Scientific Research Society, 1995.

* *Connectionist Models Summer School* Fellowship, Institute of Cognitive Sciences, Univ. of Colorado at Boulder, 1993.

* *Edna Bissette* Award, Duke University, 1993.

* *International Research Exchange (IREX)* Fellowship, 1992.

* *Soros* Travel Grant, 1992.

* Over 10 prizes, including the *2nd Prize at the National Physics Olympiad* (Bucharest, 1986), obtained at several student competitions (*Mathematics, Computer Science, Physics, Systems Theory*, and *Electricity Fundamentals*) while V. D. was with the Polytechnic Institute of Iasi, Romania.

Publications

* Dragoi, V. (in press). A Model of Contextual Interactions in Primary Visual Cortex: Examining the Influence of Corticogeniculate Feedback. In: Computational Neuroscience '96. (Bower JM, ed.). Plenum Press.

* Dragoi, V., & Staddon, J. E. R. (in review). The Dynamics of Operant Conditioning. *Psychological Review.*

* Dragoi, V. (1997). A Dynamic Theory of Acquisition and Extinction in Operant Learning. *Neural Networks*, 10, 201-229.

* Dragoi, V. (1995). Neural Dynamics of Form Perception: Geometrical Illusions and Aftereffects. *Proceedings of World Congress on Neural Networks*, 3, 275-278.

* Dragoi, V., & Staddon, J. E. R. (1993). A Competitive Neural Network Model for the Process of Recurrent Choice. In M. C. Mozer, P. Smolensky, D. S. Touretzky, J. L. Elman, & A. S. Weigend (Eds.), *Proceedings of the 1993 Connectionist Models Summer School* (pp. 65-73). Hillsdale, NJ: Erlbaum Associates.

* Dragoi, V. (1992). Structural Organization of Boolean Cellular Automata. In D. Ruck (Ed.), *Science of Artificial Neural Networks.* Proceedings of the SPIE. **1710**, 123-132.

* Dragoi, V., & Buhusi, C. (1991). A New Learning Algorithm Based on Neural Relocation. In T. Yamakawa, H. Teodorescu, M. Rascanu (Eds.), *Proceedings of the 6th International Conference on Fuzzy Systems and Artificial Intelligence IFSAI 1991*, 183-189. Iasi University Publ. House, Romania.

* Dragoi, V. (1991). Contextual Organization of Cellular Automata. In T. Yamakawa, H. Teodorescu, M. Rascanu (Eds.), *Proceedings of the 6th International Conference on Fuzzy Systems and Artificial Intelligence IFSAI 1991*, 173-182. Iasi University Publ. House, Romania.