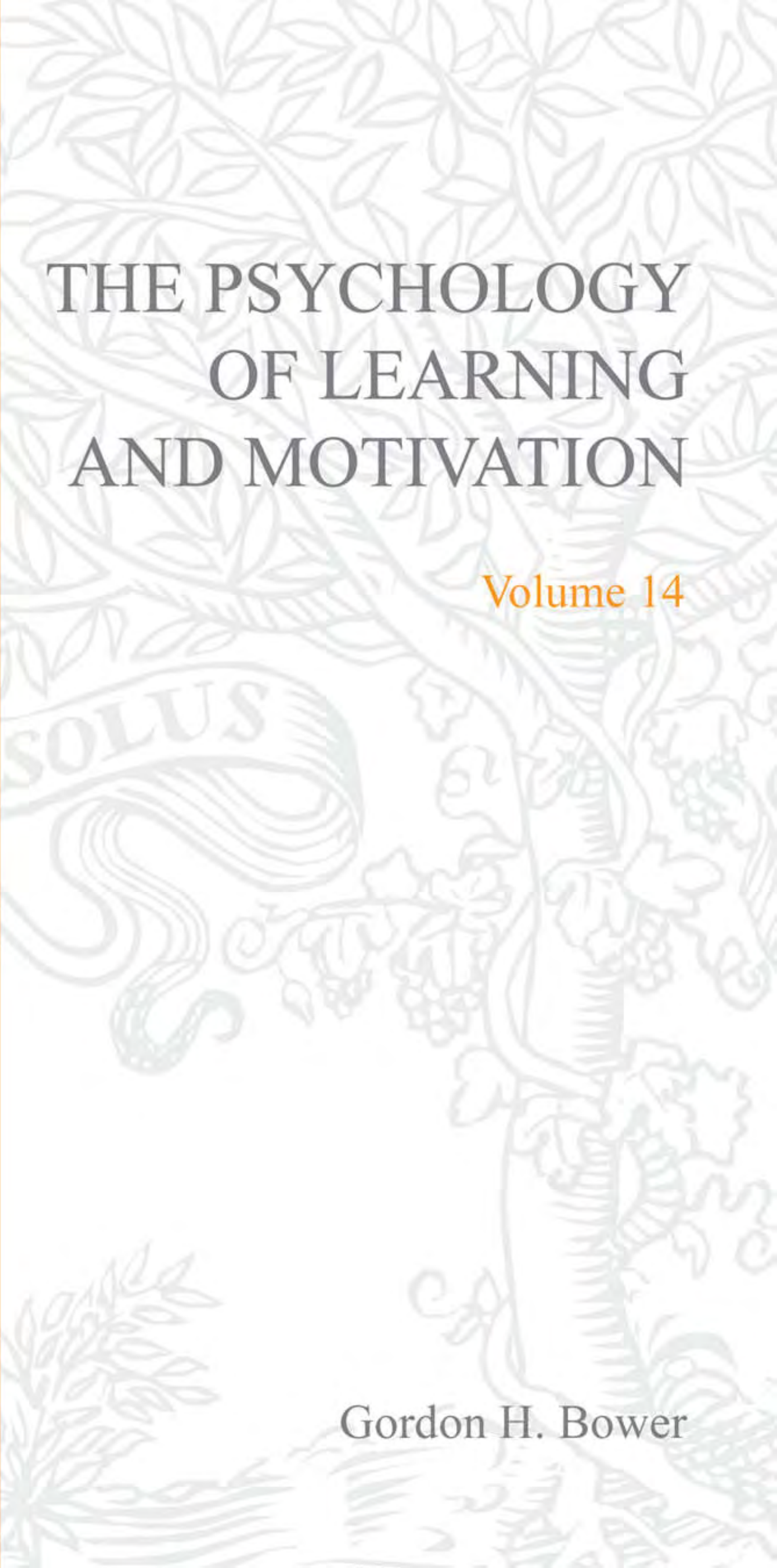


THE PSYCHOLOGY OF LEARNING AND MOTIVATION

Volume 14

Gordon H. Bower

THE PSYCHOLOGY OF LEARNING AND MOTIVATION

Advances in Research and Theory

VOLUME 14

CONTRIBUTORS TO THIS VOLUME

A. Charles Catania

Peter A. de Villiers

R. J. Herrnstein

Ronald E. Johnson

Timothy D. Johnston

Marcia Ozier

Jeroen G. W. Raaijmakers

Richard M. Shiffrin

William Timberlake

M. T. Turvey

THE PSYCHOLOGY OF LEARNING AND MOTIVATION

Advances in Research and Theory

EDITED BY GORDON H. BOWER
STANFORD UNIVERSITY, STANFORD, CALIFORNIA

Volume 14

1980

ACADEMIC PRESS

A SUBSIDIARY OF HARCOURT BRACE JOVANOVICH, PUBLISHERS

New York • London • Toronto • Sydney • San Francisco

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 66-30104

ISBN 0-12-543314-X

PRINTED IN THE UNITED STATES OF AMERICA

80 81 82 83 9 8 7 6 5 4 3 2 1

CONTENTS

A MOLAR EQUILIBRIUM THEORY OF LEARNED PERFORMANCE

William Timberlake

FISH AS A NATURAL CATEGORY FOR PEOPLE AND PIGEONS

R. J. Herrnstein and Peter A. de Villiers

FREEDOM OF CHOICE: A BEHAVIORAL ANALYSIS

A. Charles Catania

A SKETCH OF AN ECOLOGICAL METATHEORY FOR THEORIES OF LEARNING

Timothy D. Johnston and M. T. Turvey

SAM: A THEORY OF PROBABILISTIC SEARCH OF ASSOCIATIVE MEMORY

Jeroen G. W. Raaijmakers and Richard M. Shiffrin

MEMORY-BASED REHEARSAL

Ronald E. Johnson

INDIVIDUAL DIFFERENCES IN FREE RECALL: WHEN SOME PEOPLE REMEMBER BETTER THAN OTHERS

Marcia Ozier

LIST OF CONTRIBUTORS

Numbers in parentheses indicate the pages on which the authors' contributions begin.

A. Charles Catania, Department of Psychology, University of Maryland Baltimore County, Catonsville, Maryland 21228 (97)

Peter A. de Villiers, Department of Psychology, Smith College, Northampton, Massachusetts 01060 (59)

R. J. Herrnstein, Department of Psychology and Social Relations, Harvard University, Cambridge, Massachusetts 02138 (59)

Ronald E. Johnson, Psychological Sciences and Education, Purdue University, West Lafayette, Indiana 47906 (263)

Timothy D. Johnston, North Carolina Division of Mental Health Research Section, Raleigh, North Carolina 27611 (147)

Marcia Ozier, Department of Psychology, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada (309)

Jeroen G. W. Raaijmakers, Psychology Department, University of Nijmegen, Nijmegen, The Netherlands (207)

Richard M. Shiffrin, Department of Psychology, Indiana University, Bloomington, Indiana 47405 (207)

William Timberlake, Department of Psychology, Indiana University, Bloomington, Indiana 47405 (1)

M. T. Turvey, University of Connecticut, Storrs, Connecticut 06268 and Haskins Laboratories, New Haven, Connecticut 06510 (147)

This Page Intentionally Left Blank

CONTENTS OF PREVIOUS VOLUMES

Volume 1

Partial Reinforcement Effects on Vigor and Persistence
Abram Amsel
A Sequential Hypothesis of Instrumental Learning
E. J. Capaldi
Satiation and Curiosity
Harry Fowler
A Multicomponent Theory of the Memory Trace
Gordon Bower
Organization and Memory
George Mandler
Author Index—Subject Index

Volume 2

Incentive Theory and Changes in Reward
Frank A. Logan
Shift in Activity and the Concept of Persisting Tendency
David Birch
Human Memory: A Proposed System and Its Control Processes
R. C. Atkinson and R. M. Shiffrin
Mediation and Conceptual Behavior
Howard K. Kendler and Tracy S. Kendler
Author Index—Subject Index

Volume 3

Stimulus Selection and a "Modified Continuity Theory."
Allan R. Wagner
Abstraction and the Process of Recognition
Michael I. Posner
Neo-Noncontinuity Theory
Marvin Levine
Computer Stimulation of Short-Term Memory: A Component-Decay Model
Kenneth R. Laughery
Replication Processes in Human Memory and Learning
Harley A. Bernbach
Experimental Analysis of Learning to Learn
Leo Postman
Short-Term Memory in Binary Prediction by Children: Some Stochastic Information Processing Models
Richard S. Bogartz
Author Index—Subject Index

Volume 4

Volume 5

Volume 6

Volume 7

Volume 8

Edward E. Smith, Lance J. Rips, and Edward J. Shoben

Working Memory
Alan D. Baddeley and Graham Hitch

The Role of Adaptation Level in Stimulus Generalization
David R. Thomas

Recent Developments in Choice
Edmund Fantino and Douglas Navarick

Reinforcing Properties of Escape from Frustration Aroused in Various Learning Situations
Helen B. Daly

Conceptual and Neurobiological Issues in Studies of Treatments Affecting Memory Storage
James L. McGaugh and Paul E. Gold

The Logic of Memory Representations
Endel Tulving and Gordon H. Bower

Subject Index

Volume 9

Prose Processing
Lawrence T. Frase

Analysis and Synthesis of Tutorial Dialogues
Allan Collins, Eleanor H. Warnock, and Joseph J. Passafiume

On Asking People Questions about What They Are Reading
Richard C. Anderson and W. Barry Biddle

The Analysis of Sentence Production
M. F. Garrett

Coding Distinctions and Repetition Effects in Memory
Allan Paivio

Pavlovian Conditioning and Directed Movement
Eliot Hearst

A Theory of Context in Discrimination Learning
Douglas L. Medin

Subject Index

Volume 10

Some Functions of Memory in Probability Learning and Choice Behavior
W. K. Estes

Repetition and Memory
Douglas L. Hintzman

Toward a Framework for Understanding Learning
John D. Bransford and Jeffrey J. Franks

Economic Demand Theory and Psychological Studies of Choice
Howard Rachlin, Leonard Green, John H. Kagel, and Raymond C. Battalio

Self-Punitive Behavior
K. Edward Renner and Jeanne B. Tinsley

Reward Variables in Instrumental Conditioning: A Theory
Roger W. Black

Subject Index

Volume 11

Levels of Encoding and Retention of Prose
D. James Dooling and Robert E. Christiaansen

Mind your *p's* and *q's:* The Role of Content and Context in Some Uses of And, Or, and If
Samuel Fillenbaum

Encoding and Processing of Symbolic Information in Comparative Judgments
William P. Banks

Memory for Problem Solutions
Stephen K. Reed and Jeffrey A. Johnson

Hybrid Theory of Classical Conditioning
Frank A. Logan

Internal Constructions of Spatial Patterns
Lloyd R. Peterson, Leslie Rawlings, and Carolyn Cohen

Attention and Preattention
Howard Egeth

Subject Index

Volume 12

Volume 13

THE PSYCHOLOGY OF LEARNING AND MOTIVATION

Advances in Research and Theory

VOLUME 14

This Page Intentionally Left Blank

A MOLAR EQUILIBRIUM THEORY OF LEARNED PERFORMANCE[1]

William Timberlake

INDIANA UNIVERSITY
BLOOMINGTON, INDIANA

I. Introduction to the Functionalist Tradition

The behaviorist study of learning is rooted in at least two traditions: associationism, usually identified with the British Empiricists; and a more

---

[1] I thank Dan Bernstein and Mark Wozny for providing me with unpublished data, Jose Bauermeister and Academic Press for permission to adapt two figures from Bauermeister (1975), and Richard Ellis and Suzanne Hull for their fine work on the figures. Preparation of this chapter was partly supported by NSF Grant BNS 79-15117. The text profited from the criticisms of Dan Bernstein, Richard Ellis, Eliot Hearst, Ed Konarski, Connie Mueller, Phil Podsakoff, Holly Stocking, and Mark Wozny.

ISBN 0-12-543314-X

amorphous functionalist tradition, concerned with the adaptive and regulatory aspects of behavior. The associationist tradition has been concerned primarily with laws that govern the formation of stimulus and stimulus–response associations through experience. Experimenters in this tradition have emphasized the frequency and nature of pairings between events, and how these events are represented in the organism. In contrast, the functionalist tradition has focused on behavior as the mediator between the environment and the needs of the organism (Angell, 1907). Proponents of this tradition have related behavior to adaptive set points (equilibrium states)[2] such as pleasure, or the absence of pain (Digby, 1644; Spencer, 1883); maximum utility (Bentham, 1789; McFarland, 1977); optimal arousal (Berlyne, 1960; Fiske & Maddi, 1960; Walker, 1964); and homeostatic balance (Cannon, 1934; Dashiell, 1928; Richter, 1942).

Concern with associative variables has dominated the research and thinking of most learning theorists. There has been insufficient recognition that learning is one of many adaptive devices by which the organism adjusts to the challenges of its environment. The purpose of this article is to partially restore the balance between associative and functionalist analyses by focusing on two important problems in the functional tradition: the identification of circumstances that produce learned changes in behavior, and the determination of set points that regulate the amount of such change. It is not my intention in this analysis to deny the importance of associative variables, but to emphasize the regulatory framework within which they operate.

A. THORNDIKE'S CONTRIBUTION

Thorndike (1911) launched the behaviorist study of learning by combining aspects of both the associational and functional traditions within the framework of the experimental study of behavior (Hilgard, 1948). In the tradition of associationism, Thorndike investigated the development

---

[2] Set point and equilibrium state are related concepts borrowed from systems theory. They refer to stable characteristics of a functioning system which the operation of the system tends to maintain under a variety of challenges. A familiar example of a set point is the setting on the thermostat in a home heating system. The operation of the heating system maintains the temperature of the house at the thermostat setting by appropriate reactions to deviations from it. The present use of set point calls attention to the tendency of a living organism (a system) to maintain or approach characteristic states. In most cases the set point is tied closely to behavior, and the behavior of the organism will typically result in approach to that set point in the face of forced deviations from it. In the sense that a thermostat setting regulates the temperature of the house, a set point regulates behavior by controlling its direction and extent.

of the strength of associations between the stimulus situation and a response. On the functional side, Thordike emphasized the regulation of behavior with respect to a set point, the satisfying state of affairs. In a single stroke, Thorndike united the associational and functional traditions by arguing that associative bonds were formed within the context and in the service of the regulation of behavior with respect to a satisfying state of affairs.

Unfortunately, Thorndike's marriage of associational and functional traditions proved to be a sterile union for the functionalist approach. Thorndike's concern with regulation was clear in his procedures for producing a satisfying state of affairs. He simply reduced the organism's access to a common behavior, such as eating, or remaining near other animals, and then presented access to this state contingent on performing an instrumental response. However, Thorndike's formal definition of a satisfying state of affairs as "one that the organism does nothing to avoid, often doing things which maintain or renew it," was less clear (Thorndike, 1913, p. 2). Without further guidance it was difficult for investigators to agree on the same measures of nonavoidance and maintenance. Thus, in practice, the set points regulating behavior were not specified.

Because of Thorndike's vagueness and the reluctance of behaviorists to embrace the mentalistic overtones of a satisfying state of affairs, most researchers focused on the associational mechanisms underlying learned performance. Some theorists, encouraged by the discoveries of Pavlov (1928), attempted to reduce the regulatory (striving) aspects of behavior to a secondary principle, derivable from the primary principles of conditioning. Hull (1935) in his review of Thorndike's book, *The Fundamentals of Learning,* commented:

> There is implicit in Thorndike's formulation not only a law of learning but, in addition, an hypothesis correlating learning with motiviation. . . . This serves to raise the fundamental question as to which of the two variables is primary. Does the motivation (striving) produce the learning (strengthening), or does the learning produce the motivation, or does some third and still more basic process produce both? Thorndike seems to imply that the striving is primary, though he does not specifically say this. The reviewer, on the other hand, rather inclines to the view that the "conditioning" or strengthening is primary. By this is meant that striving can probably be derived as a theorem from the principles of conditioning as basic assumptions. (Hull, 1935, p. 821)

From the perspective of the present, Hull's either/or question is not appropriate. Both associative and regulatory processes, however intertwined, are likely to exist. Since in phylogeny the ability to regulate precedes the ability to learn, it is unlikely that the regulatory processes of an organism would depend exclusively on learning. Hull (1943) subsequently altered his view to postulate the existence of basic tissue needs in the organism that could not be reduced to associative factors. Other theoreticians remained on Hull's original tack, primarily emphasizing associative rather than regulatory variables (e.g., Estes, 1950; Guthrie, 1935).

During this time, regulatory concepts received considerable attention from biological theorists such as Craig (1918), Loeb (1918), and Jennings (1906), and psychologists such as Carr (1925), Dashiell (1928), Warden, Jenkins, and Warner (1936), and Lewin (1942). These researchers emphasized that behavior was organized around physiological set points and equilibrium states. For example, Carr (1925) hypothesized "When an organism needs food or water, there is a persistent internal stimulus that produces activity until the need is satisfied."

Interest in regulatory processes markedly increased when medical research provided physiological evidence for their existence. Following the earlier work of Bernard (1878/1966), Cannon (1918,1934) introduced the concept of homeostasis to describe the many precise adjustments that served to regulate physiological equilibria within the body. Though Cannon's emphasis was on internal mechanisms, he explicitly considered the role of behavior in maintaining internal homeostatic balance. During this period, Richter (1922, 1927, 1942) began his studies demonstrating the correlation between general activity, diet selection, and internal deficits produced by deprivation or surgery.

B. DRIVE THEORY

The majority of learning researchers quickly assimilated the concept of tissue deficits and homeostasis to the already existing notion of drive (Woodworth, 1918) and used the concept of drive reduction in much the way Thorndike treated a satisfying state of affairs. On the surface, drive theory seemed ideally suited to restore regulatory processes to a central role in learned behavior. Since drives were identified with physiological imbalances (tissue needs), a satisfying state of affairs could be defined by the intake of substances that reduced those imbalances. The set point of behavior was determined indirectly by the amount of a substance required to restore the internal balance.

However, in practice, drive theory proved inadequate in identifying either the dimensions of regulation or the set points that determined amount of behavior. Learning clearly occurred without reduction of tissue needs (Berlyne, 1960; Harlow, Harlow, & Meyer, 1950; Sheffield, Wulff, & Backer, 1951) and did not necessarily occur when tissue needs were reduced (Richter, 1942). If regulation provided the context and impetus for learning, something other than tissue needs was regulated.

In the absence of specific physiological or behavioral set points, researchers predicted performance on the assumption that drive increased monotonically with hours of deprivation (e.g., Hull, 1943). But, such monotonicity, when demonstrated, did not necessarily last past 24 hr (Birch, Bernstein, & Clark, 1958; Dufort & Wright, 1962). Further, the precise relation between deprivation and behavior varied with how the drive was measured (Haverland, 1954; Miller, 1955), what incentive items were employed (Elliott, 1928), and what species were observed (Campbell, Smith, Misanin, & Jaynes, 1966). Some incentive items such as sexual behavior and activity appeared unrelated to tissue needs, and did not show monotonic relations with hours of deprivation.

Recent researchers have questioned whether an organism would long survive if its behavior were regulated primarily by tissue needs. Such an organism would always be on emergency status and close to death. Instead, as Fitzsimmons (1972) and Collier, Hirsch, and Hamlin (1972) pointed out, organisms regulate drinking and eating largely in apparent anticipation of future needs. Similarly, organisms reliably regulate breathing in the absence of a current oxygen deficit.

In short, despite its apparent potential, drive theory specified neither the dimension nor set points of regulation. Though drive was operationally defined by hours of deprivation, the accuracy of this definition is questionable. Last, the reduction of tissue needs is neither the necessary nor sufficient condition for producing or regulating learned behavior.

C. THE EMPIRICAL APPROACH

As the conceptual underpinnings of drive theory failed, experimenters assumed an empirical stance that focused on repeatability and prediction of results in familiar paradigms. In place of set points, theorists used deprivation states with predictable effects and equations with parameters that were fit from the data (e.g., Hull, 1943, and more recently, Herrnstein, 1970). Many researchers attempted to minimize any effect of regulatory variables by using brief experimental sessions, highly deprived organisms, and small amounts of reward.

As the identification of set points became dependent on inference from learned performance, so did the identification of a satisfying state of affairs. By using familiar circumstances and species, experimenters were able to identify reinforcers based on their past effects on behavior. Meehl (1950) elevated this seat-of-the-pants technique to the status of a general law with the claim that reinforcers are transituational in their effects. As a description of the basic assumption that experimenters used in research, Meehl's empirical law of effect was accurate. As a formal law, it was a failure because it did not specify enough of the critical circumstances that produced reinforcement (Premack, 1965, 1971; Timberlake & Allison, 1974). In some sense (and with tongue in cheek), applying the empirical law of effect can be compared to operating a light switch. If the switch works, things are clear; if it does not, you're left in the dark.

In summary, the "empirical" approach to specifying reinforcers and set points added little to our knowledge of the regulatory determinants of learned performance. At a conceptual level, the empirical law of effect was a temporary repair of learning theory. At an empirical level, the empirical law of effect described a pragmatic approach that worked, but within poorly defined limits. To advance our understanding of learned performance, it is important to systematize successes and failures in the transituational identification of reinforcers.

D. PREMACK'S APPROACH

The functionalist approach to learned behavior owes much to the research and theorizing of David Premack (1959, 1965, 1971). Though Premack rarely emphasized the regulatory aspects of his work, he provided a means of identifying reinforcers (the probability-differential hypothesis), a technique for assessing the set points regulating behavior (the paired baseline), and evidence for the importance of molar equilibrium states in controlling learned behavior.

Premack differed from most other empirically oriented theorists in that he treated a reinforcer as a part of behavior, a combination of the stimulus presented and the behavior associated with the stimulus (Premack, 1971). For clarity here, I will refer to the stimulus aspect of a reinforcer as the reinforcing stimulus, and to the response aspect as the contingent response (Timberlake & Allison, 1974).

As a replacement for the empirical law of effect, Premack (1965, p. 131) formulated the probability-differential hypothesis: "For any pair of responses, the more probable one will reinforce the less probable one," (but the reverse will not occur). Premack determined the relative probability of two responses by measuring their durations in a paired baseline. The paired baseline was identical to the contingency session, except that the stimuli controlling both responses were freely and simultaneously available.

The paired baseline also served to identify the set points regulating learned behavior. In essence, measures of responding in the paired baseline provided an empirical substitute for the inference of set points from drive theory.[3] Premack (1959, 1965) used the baseline of the contingent response as an implicit set point for behavior when he predicted that the increase in responding under a schedule would be directly related to the probability-differential between the contingent and instrumental responses (see also Donahoe, 1977).

Premack's approach was clearly superior to the previous ''empirical'' approach in identifying the circumstances for reinforcement before the schedule was imposed and in predicting the amount of subsequent change in responding. However, some problems remained. A potential problem with the paired baseline procedure was that not all reinforcing stimuli are associated with an easily measured response (Hilgard & Bower, 1975). This objection has been met by using a simple indicator response to assess the probability of being in a particular response state. For example, Timberlake and Allison (1974) showed that the baseline of avoiding shock could be measured by allowing a rat to hold a bar in baseline in order to remain in the shock-free state. In the contingency, the package of bar-holding-for-shock-free-state was made contingent on an instrumental response.

Premack (1965) himself discovered the major shortcoming of his approach in his analysis of the role of the schedule in producing changes in instrumental responding. In an intriguing study, Premack (1965) arranged a schedule relating the lower probability response of wheel running and the higher probability contingent response of drinking in such a way that by the time the subject had run its baseline amount it also had drunk its baseline amount. He accomplished this by imposing a schedule with the same ratio of instrumental running to contingent drinking that was present in the baseline. Under these special schedule conditions, there was no change in responding. Thus, the probability-differential condition was not sufficient to produce an increase in instrumental responding.

At this point Premack had all the pieces necessary to assemble an equilibrium theory that would integrate and extend the regulatory analysis of learned behavior. He knew that the probability-differential hypothesis was incomplete because a contingent response of higher probability had

---

[3] It is interesting that Skinner (1938) used a similar baseline technique in validating barpressing as a measure of drive. Skinner (1938) first measured how much a rat ate in a free baseline by recording the number of times the rat pushed open the door to the food tray when each push produced a pellet. Later, under a fixed-ratio one schedule, the procurement of the pellet was made contingent on pressing a bar. The cumulative record for barpressing appeared to follow closely the time course of the original feeding response. Skinner concluded, ''In measuring the strength of a drive, we are in reality only measuring the strength of behavior.'' Unfortunately, Skinner did not develop this insight further, and subsequent experimenters abandoned the approach.

no effect when the scheduled ratio of instrumental and contingent responding equaled the baseline ratio. He also knew that the probability-differential hypothesis was accurate when the scheduled ratio differed from the baseline ratio. As we will see, it is a small step from these facts to the disequilibrium hypothesis that a ratio schedule affects behavior only if the ratio of responding it specifies differs from the baseline ratio of responding.

Unfortunately, instead of focusing on the a priori difference between schedule and baseline ratios, Premack focused on the actual reduction in contingent responding that occurred under most schedules that increased the instrumental response. He claimed that reinforcement could not be initiated in the absence of a schedule that produced such a reduction (Premack, 1965). Though an actual reduction in contingent responding is often correlated with the presence of an a priori difference between the schedule and baseline response ratios, the a priori difference concept has several advantages over the reduction approach. First, in the difference approach the presence and direction of the difference can be determined ahead of time and used to predict whether and in what way a schedule will change instrumental responding. On this point Premack's actual reduction approach has no apparent advantage over the empirical law of effect. In both cases, one must wait until the data are in to make an accurate prediction. Second, there is reasonable evidence that some schedules that increase instrumental responding produce very little, if any, reduction in contingent responding. For example, Premack (1959) reported that his Cebus monkeys sometimes increased instrumental responding sufficiently to maintain the contingent response at its baseline level. Wasik (1969), using humans, and Harrison and Schaeffer (1975), using rats, reported significant increases in instrumental responding in the absence of significant decreases in the contingent response. These findings cast considerable doubt on the importance of an actual reduction in contingent responding.

In summary, Premack provided a reasonable and useful means of identifying reinforcers and assessing the set points that regulated behavior. In analyzing the role of the schedule in changing responding, he focused on the actual reduction in contingent responding rather than the a priori difference between the schedule and baseline ratios. In his last article, Premack (1971) attempted to integrate the actual reduction in the contingent response with the concept of momentary probability to explain the causal status of the reduction. I will comment on this attempt later.

E. THE EQUILIBRIUM APPROACH

The equilibrium viewpoint latent in Premack's approach apparently was developed independently by Eisenberger, Karpman, and Trattner

(1967), Timberlake and Allison (1974), and Marmaroff (1971) reported in Dunham (1977). In simplest form, an equilibrium theory proposes that there exists a stable set of conditions that an organism will approach or maintain under circumstances that perturb or challenge these conditions. In the present approach, the stable set of conditions is assumed to be expressed in the behavior of an organism in a free baseline when the stimuli controlling all responses are freely and simultaneously available. If a schedule perturbs the equilibrium condition by forcing responding away from its baseline expression, the organism is presumed to act to reduce the resultant disequilibrium (within limits set by its past and present experience and evolutionary history).

For purposes of illustration, Table I shows a simple comparison between the empirical law of effect and the equilibrium approach with respect to the causal circumstances, setting conditions, and results of a schedule that produces learned performance. The equilibrium approach specifies the causal circumstances in a priori fashion, while the empirical law of effect depends on prior demonstration of efficacy. Further, the setting conditions specified by the empirical law of effect are more complex. Both approaches require that conditions for learning be present, but the empirical law of effect also requires that the organism be deprived of the contingent response outside the experimental session (Meehl, 1950).

TABLE I

TWO APPROACHES TO LEARNED PERFORMANCE

Theory	Level	Causal circumstance	Setting conditions	Result
Empirical law of effect	Conceptual	Strengthening or punishing agent	1. Drive or incentive state 2. Learning possible	Strengthen or weaken response in stimulus conditions
	Empirical	Present previously established reinforcing or punishing stimulus	1. Known deprivation schedule 2. Learning conditions	Increase or decrease instrumental response
Equilibrium theory	Conceptual	Disequilibrium condition	1. Reliable equilibrium 2. Learning possible	Reduce deviations from equilibrium
	Empirical	Impose schedule that disrupts baseline relations	1. Stable baselines 2. Learning conditions	Responses reapproach baseline levels

In contrast, the equilibrium approach can specify conditions under which a deprived animal will show no increase in instrumental responding, other conditions under which an undeprived animal will show an increase in instrumental responding, and other conditions in which the reverse of these effects should occur. Last, the predictions of equilibrium theory are more precise, because the baseline set points place limits on responding. The empirical law of effect places no limits on changes in responding.

The present equilibrium approach is a more complete realization of the response deprivation approach of Timberlake and Allison (1974). Response deprivation refers to the disequilibrium condition presumed to underlie increases in instrumental responding. It is defined as occurring if the subject, by performing its baseline level of the instrumental response, is unable to perform the contingent response at its baseline level. The organism is presumed to adapt to the initial disequilibrium between the schedule requirement and the baseline by increasing instrumental responding.

Heth and Warren (1978) made explicit that there are two sorts of disequilibrium conditions by defining response satiation as the condition in which the subject, by performing its baseline level of the instrumental response, is forced to exceed its baseline level of the contingent response. Under these circumstances, the subject is presumed to adapt to the initial disequilibrium by decreasing instrumental responding.

The remainder of this article attempts to develop and test further the molar equilibrium approach that underlies the notions of response deprivation and satiation. In Section II, I will present and discuss a more formal representation and statement of the approach. In Sections III and IV, I will review a variety of predictions and data bearing on this approach, and briefly consider the importance of the molecular determinants of learned performance.

II. Toward a Molar Equilibrium Theory of Learned Performance

Equilibrium theories in many areas of science lend themselves to formal representation. In this section I will present an intuitive graphic representation, followed by a more formal statement and discussion of the general assumptions underlying the graphic representation.[4]

---

[4] I thank John Staddon and Will Vaughan for their patience in introducing me to this form of representation. They cannot be held accountable for the use to which I put it.

A. REPRESENTATION

The essential components of the graphic approach are illustrated in Fig. 1 for a fixed-ratio schedule relating two responses. A ratio schedule was used because most of the data reported in this paper are from such schedules. Each panel in Fig. 1 represents a two-dimensional behavior space in which each dimension corresponds to a response. The paired baseline of the contingent and instrumental responses is represented by the point (O_c, O_i). Possible behavior on a fixed ratio schedule corresponds to a collection of points falling on a straight line through the origin, with a positive slope. The slope of the line corresponds to the ratio of the instrumental requirement to the contingent payoff (I/C).[5] The diagonal line in Fig. 1a represents a schedule in which the ratio of the instrumental requirement to the contingent payoff (I/C) is greater than the ratio of the baselines of instrumental and contingent responding (O_i/O_c).

The difference between the baseline and scheduled ratios constrains behavior such that the subject cannot regain its baseline level of both responses at once. In Fig. 1a, if the subject were to respond at its baseline level of the instrumental response (the thin horizontal line), it would fall short of its baseline level of the contingent response by amount D_c (the contingent deficit condition). On the other hand, if the subject were to respond at its baseline level of the contingent response, it would exceed its baseline level of the instrumental response by amount E_i (the instrumental excess condition). How the subject will change its behavior can be predicted by determining what movement on the schedule line reduces the disequilibrium imposed by the schedule. In Fig. 1a, increasing the instrumental response above its baseline level will decrease the contingent deficit condition, while decreasing the contingent response below its

---

[5] There are several important qualifications of this representation. First, responding will not fall exactly on the schedule line unless the ratio schedule is reciprocal (Mazur, 1975; Timberlake & Allison, 1974), and the session ends after the subject has completed a particular sequence of the instrumental and contingent responses without beginning a new one. A reciprocal schedule guarantees the scheduled ratio between instrumental and contingent responding by allowing access to each response only after the subject has performed the prescribed amount on the other. Schedules with unlimited hold on access to the contingent response will often approximate the schedule line, but with instrumental responses of high baseline there may be overshooting of the instrumental requirement (Schaeffer, 1966). Schedules with limited (timed) hold on access to the contingent response will always produce a higher ratio than the one stated for a deficit schedule. Limited hold schedules also are of no use in conditions of excess for the contingent response.

The second qualification of this representation is that the schedule term and baseline amount of each response must be measured in the same units. However, it is not necessary for both responses to be measured in the same units.

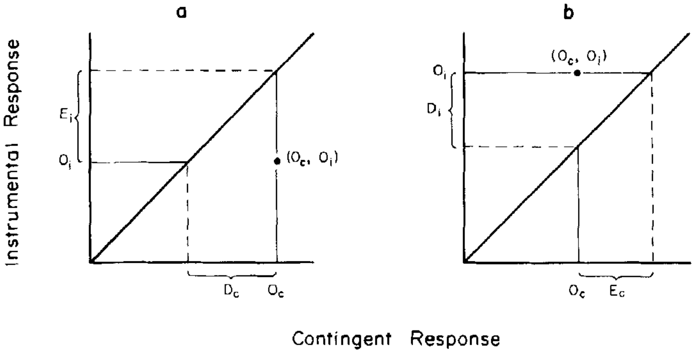


Fig. 1. Disequilibrium conditions in a two-dimensional behavior space, in which each dimension represents the possible values of one response. Disequilibrium conditions are produced by a difference between the scheduled and baseline ratios of the two responses. (a) A combination of baseline point (O_c, O_i) and fixed-ratio schedule (the 45 degree line) that produces a deficit condition for the contingent response (D_c) and an excess condition for the instrumental response (E_i). Responding of the subject should fall near the schedule line and between the points at which a line from O_i and a line from O_c intersect the schedule line. (b) A combination of baseline and schedule values that produces an excess condition for the contingent response (E_c) and a deficit condition for the instrumental response (D_i). Again responding should fall near the schedule line and between the points at which lines from the two baseline values intersect the schedule line.

baseline will decrease the instrumental excess. If subjects respond to decrease both disequilibrium conditions, they should increase the instrumental response above baseline, but rarely enough to regain the baseline of the contingent response.

Figure 1b is analogous to Fig. 1a except the contingent response has a very low baseline level, and the scheduled ratio is less than the ratio of the instrumental and contingent responses in baseline($I/C < O_i/O_c$). In this case, if the subject responds at its operant level of the instrumental response, it will perform the contingent response at more than its baseline level. On the other hand, if the subject performs its baseline level of the contingent response, it would be necessary to perform less than its baseline level of the instrumental response. If subjects respond to decrease both disequilibrium conditions, they will decrease instrumental responding and increase contingent responding with respect to baseline.

The reader will recognize the first example as a typical appetitive instrumental schedule in which the subject's normal behavior results in a deficit condition for the contingent response; the second example is typical of punishment schedules in which the subject's baseline behavior produces an excess condition for the contingent response (or contingent state, if no response is specified). It can be seen in the figure that each schedule actually produces the same two results, an excess of one response if the other is performed at its baseline level, and a deficit of the second if the first is performed at its baseline level. Thus, from an equilibrium viewpoint, there is no necessary distinction between a punishment and a reward schedule. In one case the focus is on the effects of the forced deficit on behavior, and in the other the focus is on the effects of the forced excess.[6]

In the remainder of the article I will follow tradition and emphasize changes in instrumental responding as a function of disequilibrium conditions involving the contingent response. However, the above analysis should make clear that a change in contingent responding will also be characteristic of behavior under schedules. As we will see, a complete equilibrium theory will need to deal with disequilibrium conditions involving both responses. I will also limit my review to ratio schedules relating two responses, though the graphic approach can be generalized to three or more responses (Rachlin & Burkhard, 1978; Staddon, 1979), and to any schedule that can be represented in the present behavior space (Staddon, 1979).

Figure 2 shows a unique extension of the graphic approach to three responses: instrumental, contingent, and background responses (Rachlin & Burkhard, 1978; Staddon, 1979). These three responses can be represented in a "temporal triangle" with one response at each vertex. The vertices actually fall on the axes of a three-dimensional space, and the triangle represents those points in the space for which the durations of the responses sum to the total session time. In this figure, a ratio schedule is represented by the diagonal line rising from right to left. It is the projection on this plane of the diagonal schedule line shown in Fig. 1a. It can be seen that if the subject performs the instrumental response at its baseline, indicated by the line parallel to the bottom of the triangle, a contingent deficit results. By extension, a representation of contingent excess can also be obtained.

B. GENERAL ASSUMPTIONS

In this section I present the general assumptions underlying the present equilibrium approach.

---

[6] Measuring the disequilibrium in terms of amount of deficit and/or excess is somewhat arbitrary. The disequilibrium could also be measured by the difference in the angle between a line drawn through the baseline point and the schedule line, or by the distance from the baseline point to the schedule line along a perpendicular. As far as I can tell, all of these techniques make the same predictions concerning the presence and direction of change in instrumental responding, and similar predictions for the amount of change. A major difference is that the last two measurement techniques do not distinguish the disequilibrium conditions for the instrumental and contingent responses (though these amounts could be calculated).

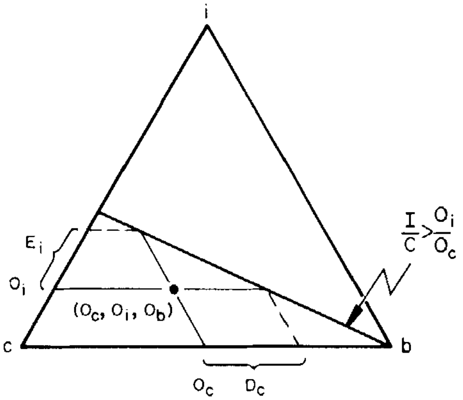


Fig. 2. A "temporal triangle" in which every point is a combination of instrumental, contingent, and background response durations that totals the session time. The vertices of the triangle fall on the axes of a three-dimensional behavior space. The combination of schedule line and baseline point (O_c, O_i, O_b) was chosen to produce a condition of deficit for the contingent response and excess for the instrumental response.

1. To the extent that characteristics of responding in a free baseline are stable and can be recovered following schedule manipulations, there exists one or more equilibrium states that reliably instigate and control responding.

2. If a schedule forces a change in the relations among response characteristics in the absence of a change in the equilibrium states instigating and controlling them, a disequilibrium condition will be produced that the behavior of the organism will tend to reduce.

3. A disequilibrium condition will result in learned performance to the extent that responses and/or predictive stimuli are reliably associated with reduction in that disequilibrium. The reliable association may be based on present and past experience, and evolutionary history.

4. The size of the change in instrumental responding under a disequilibrium condition will be directly related to the size of the initial disequilibrium.

C. DISCUSSION

The heart of Assumption 1 is the presumption that an equilibrium state underlies any reliable characteristic of responding, no matter how complex or simple. In most of the present examples this state is measured by the duration or frequency of the instrumental and contingent responses. It might also be measured by more molecular response characteristics (Dunham, 1977; Premack, 1971), or in the context of a larger number of responses (Allison, Miller, & Wozny, 1979; Rachlin & Burkhard, 1978; Staddon, 1979).

A problem with this assumption is to determine when a reliable level of behavior is achieved. In studies of animals responding is often measured over 20–30 days before stability is produced. This procedure must be contrasted with some studies of humans in which baselines of 5 min were presumed to reflect reliable levels of responding (e.g., Klajner, 1975). Shorter baselines may be reasonable in studies of humans because humans may require less time to adjust to the situation, and may be able to interrogate themselves concerning their preferences. Bernstein (1979) reported high correlations between estimated attractiveness of behavior and subsequent time engaged in that behavior. It may also be that the reactions to novelty of many species are strong enough that in brief sessions there is stable responding across baseline and contingency periods. However, given its importance, some care must be exerted to ensure the reliable instigation of responding.

A related problem concerns the recovery of baselines following schedule conditions. Though it is unnecessary for baselines to be precisely recoverable to apply the equilibrium approach, recoverability is a potential indicator of the existence of an underlying equilibrium state. In most cases baselines are recovered consistently (e.g., Bauermeister, 1975). In some cases there is a trend over successive baselines that appears related to changes in the organism rather than to effects of the schedule conditions (Timberlake & Wozny, 1979). Though the situations are different, the contrast of this baseline recovery with the marked decrease in instrumental responding reported in some human studies (Lepper & Greene, 1978) raises intriguing and unanswered questions.

An important issue that has received insufficient attention is whether measurement of a baseline alters how the animal responds under a schedule. The tentative answer is that it does not. McIntosh (1974) found that rats nose-poking for access to a saccharin solution took approximately as many total days to reach asymptotic licking, whether a baseline exposure was included or not. McIntosh's results support the idea that the baseline allows fear and escape responses to adapt, while exploration serves to identify the response potentials of the situation prior to imposition of the schedule. I have sparse pilot data on licking schedules that support a similar conclusion. However, this conclusion must be treated cautiously without more data.

The last important issue raised by Assumption 1 is to what extent the underlying equilibrium states are adequately expressed in the baseline measures. It can be argued intuitively that the correspondence between underlying states and baseline measures will never be perfect, but it will be better the fewer the encumbrances on responding. Thus, measuring a baseline of eating by the number of barpresses on a fixed-ratio 100 schedule with an 8-ounce bar will not be as clear an estimate of the set point for eating alone as measuring the baseline number of pellets consumed when freely available. It seems reasonable that equilibrium states have limited potential to resist perturbation. Thus, the more encumbered a particular response, the less likely its baseline can be used as an effective estimate of the underlying set point.

Assumption 2, that organisms will act to reduce the disequilibrium state, has two key aspects. The first is that the factors instigating and controlling behavior should not change between the baseline state and the schedule. In practice, this means that no responses or stimuli are introduced or removed between the baseline and schedule conditions, that the schedule sessions occupy the same time as the baseline, and that the circumstances outside the experimental session remain unchanged. Some alterations produced by the schedule appear unavoidable, such as changes in instigation due to periodic presentation of access to the contingent response (Falk, 1971). Further, the mere presence of a restriction on responding may instigate particular sets of behavior apparently designed to reduce or avoid experimenter control. Kavanau (1969) reported that canyon mice would bar press to turn on a motorized wheel that the experimenter turned off, or press to turn off a motorized wheel that the experimenter turned on. Bernstein and Dearborn (1978) suggested that humans worked under schedules to place themselves in a state of potential access to the contingent response, rather than performing the contingent response immediately when available. Most such changes in instigation should not markedly alter the qualitative predictions of the equilibrium approach, but may vary the amount of change in behavior. Accurate prediction of responding under a schedule may require a more sophisticated assessment of instigation than is provided by the simple paired baseline.

The second key aspect of Assumption 2 is the definition and measurement of the disequilibrium condition. In general, a condition of disequilibrium can be discovered by inspection of the relation between the baseline and the terms imposed by the schedule (Timberlake & Allison, 1974; this article). If the circumstances prevent the subject from maintaining its baseline characteristics of responding, a disequilibrium condition occurs. In a fixed ratio schedule there are two sorts of disequilibria. A condition of deficit occurs if the subject, by performing its operant level of the instrumental response, is unable to regain its baseline level of the contingent response. A condition of excess occurs if the subject, by performing its baseline level of the instrumental response, is forced to exceed its baseline level of the contingent response. I chose the word deficit rather than deprivation to emphasize that the condition is defined before the schedule is imposed. It does not refer to the subject's actual performance under the schedule.

Assumption 3 attempts to make contact with the traditional reinforcement approach to schedule performance. Nothing in the first two assumptions suggests that the organism will necessarily show learned performance as a result of a disequilibrium condition. Assumption 3 acknowledges the importance of associative mechanisms in producing learned performance. The probability of learned performance will be enhanced by limiting response alternatives, making the instrumental response the most obvious and efficient means for reducing the disequilibrium condition, and by eliciting behavior that is phylogentically programmed to reduce the disequilibrium under natural conditions. An example of this last point is that the pairing of a key light with the presentation of food to pigeons elicits pecking, which under more natural circumstances would be instrumental in obtaining food (Hearst & Jenkins, 1974).

If the experiment produces learned behavior in the disequilibrium condition, there is the problem of distinguishing changes in behavior due to learning (the contingent effects of the schedule) from changes due the noncontingent effects of the schedule (Staddon, 1979; Timberlake, 1979). Particularly in the case of highly probable instrumental responses, there is the possibility that the reduction in contingent responding under the schedule will produce an increase in instrumental responding due merely to increased opportunity for its expression, or differential substitution for the contingent response (Timberlake, 1979). Further, intermittent access to the contingent response may produce effects on instrumental responding that range from facilitation of an instrumental response (Powell & Curley, 1978) to competition with it (Iverson, 1976). In most of these situations, neither the single nor paired baselines adequately distinguishes the nature of changes in instrumental responding.

Timberlake (1979) suggested three possible ways to separate the contingent and noncontingent effects of a disequilibrium: (*a*) Estimation of the noncontingent effects of a schedule using a model of noncontingent effects, such as that of Luce (1959)—see Timberlake (1979) and Podsakoff (1980). (*b*) Use of a control group yoked to the same pattern and amount of reward access as the experimental animals. (*c*) Use of a within-subject's baseline procedure that yokes contingent responding either to the reduced contingent responding under the schedule (massed baseline) or to the reduction in and pattern of presentation of the contingent response under the schedule (matched baseline). Dunham (1977) has considered some of these same issues, as have Bernstein and Ebbesen (1978) who suggested a random matched control in which pattern of presentation was random rather than matched to the access points during the contingency. A last control condition, the equilibrium schedule (Bernstein & Hinkle, 1977), may serve to measure the instigating effects of periodic presentation of the contigent response in the absence of a molar disequilibrium. Of these procedures, several studies have indicated that the within-subject yoked baselines are most generally useful and accurate (Timberlake, 1979; Timberlake & Wozny, 1979; though see Wozny, 1979).

Assumptions 1 and 2 identified a condition of disequilibrium as critical in producing a change in instrumental responding under a schedule (Heth & Warren, 1978; Timberlake & Allison, 1974). Assumption 3 clarified the relation of this change to learned behavior. Assumption 4 predicts the relative change in instrumental responding as a function of the size of the initial disequilibrium condition. Assumption 4 is patterned after the response deprivation hypothesis of Timberlake and Allison (1974) and is intended only as a working hypothesis, enabling the prediction of functional relations among variables.

There is a trend in this area of research toward rapid quantification (see Timberlake & Wozny, 1979). In this article I have avoided quantification to keep from obscuring important empirical and conceptual issues by a concern with statistics and fitting parameters. As we will see, the prediction of functional relations from Assumption 4 allows quite strong tests of the equilibrium approach. It may well be that some recent models are preferable to others in fitting these results, but I will leave that argument for another time.

However, the reader should be aware of a number of weaknesses in Assumption 4. First, it is of little use in predicting the choices of an animal given two or more alternative schedules. A choice assumption is clearly necessary (see Staddon, 1979, for an integrated account). Second, the size of the change in instrumental responding will depend on the schedule ratio, as well as on the initial disequilibrium. For example, if two different schedules have the same initial contingent deficit, and the same eventual reduction in the contingent response, the increase in instrumental responding will be greater for the schedule specifying the larger ratio of instrumental to contingent responding. This problem should occur only when comparing results of schedules using different responses. In Section III, I will point out more general difficulties with this assumption that may affect responding on all schedules.

III. Predictions from a Molar Equilibrium Theory

This section reviews three classes of prediction from a molar equilibrium approach: (*a*) Predictions compatible with the probability-differential approach of Premack (1965). (*b*) Predictions of response change specific to the equilibrium approach. (*c*) Functional predictions of the equilibrium approach.

A. PROBABILITY-DIFFERENTIAL PREDICTIONS

In theory, all but one of the predictions of the probability-differential approach are a special subset of predictions from a molar equilibrium theory. As will be seen, the subset is formed by combining the baseline probability conditions specified by Premack with the schedule conditions required by molar equilibrium theory. In examining the literature, it is not always possible to determine if the schedule conditions required by equilibrium theory are satisfied. However, in so far as it is possible to check, the data are compatible with the equilibrium approach. The one contradictory prediction, that instrumental responding should increase as a direct function of its baseline, also contradicts the probability-differential hypothesis that relates increased instrumental responding to the size of the probability differential (Donahoe, 1977; Premack, 1965). With the exception of this hypothesis, I will not review empirical evidence for Premack's predictions. For empirical support the reader is referred to Bernstein and Ebbesen (1978), Danaher (1974), Premack (1959,1965,1971), and Terhune (1978).

The probability-differential approach contains at least nine predictions. The first seven predict only whether there will be a change in the instrumental response. The last two predictions are functional, relating the amount of change in responding to the baselines of the instrumental and contingent responses.

1. A contingent response of higher baseline probability will reinforce an instrumental response of lower baseline probability.

The basis for this prediction is shown in Fig. 1a. If the ratio value of the schedule exceeds the ratio value of the baselines, a deficit condition will occur, and the instrumental response will increase. Since the majority of Premack's experiments used schedules with a ratio approximating one unit of instrumental requirement for one unit of payoff, the scheduled ratio should always exceed the baseline ratio when the contingent response is of higher probability.

2. A contingent response of lower baseline probability will not reinforce an instrumental response of higher baseline probability.

Figure 1b shows that with a similar one-to-one schedule and a baseline ratio greater than one, the result is a condition of excess. Thus, the instrumental response should not increase.

3. A contingent response of lower baseline probability that is forced on the subject following an instrumental response of higher baseline probability should punish (decrease) the instrumental response.

The only difference between this prediction and the previous one is that in this case the subject is forced to engage in the contingent response (adhere to the schedule line), rather than being allowed to let the access to the contingent response pass without using it. The result is a condition of forced excess that should produce a decrease in the instrumental response. If the schedule allowed the subject to avoid engaging in the lower probability contingent response, the effective schedule chosen by the subject might well produce no excess condition for the contingent response.

4. No reinforcement should occur if the baseline probabilities of the instrumental and contingent responses are equal.

The basis for this prediction can be seen by imagining that the baseline point falls on the one-to-one schedule line in Fig. 1a. Under this condition, no condition of deficit or excess occurs, so the animal should change neither response.

5. For any set of responses, there exists a hierarchy of reinforcement values corresponding to the hierarchy of their baseline probabilities.

The prediction of a hierarchy can be seen in Fig. 3. Figure 3a shows that if the most probable of three responses is the contingent response, the schedule produces a deficit condition with both instrumental responses, and, therefore, predicts an increase in both. Figure 3b shows that if the contingent response is the least probable response, then the schedule produces an excess condition with both instrumental responses, and, therefore, predicts no change or a decrease in instrumental responding (depending on the type of schedule). Figure 3c shows that if the response of intermediate probability is the contingent response, the schedule produces a deficit condition with the low-probability instrumental response and an excess condition with the higher probability instrumental response. Therefore, it predicts no change or a decrease in the higher probability instrumental response (depending on the type of schedule), and an increase in the lower probability instrumental response.

6. The reinforcement relation between two responses can be reversed by reversing their relative baseline probabilities.

The basis for this prediction can be seen in Fig. 4 for the responses of drinking and running. In Fig. 4a, where running is the more probable

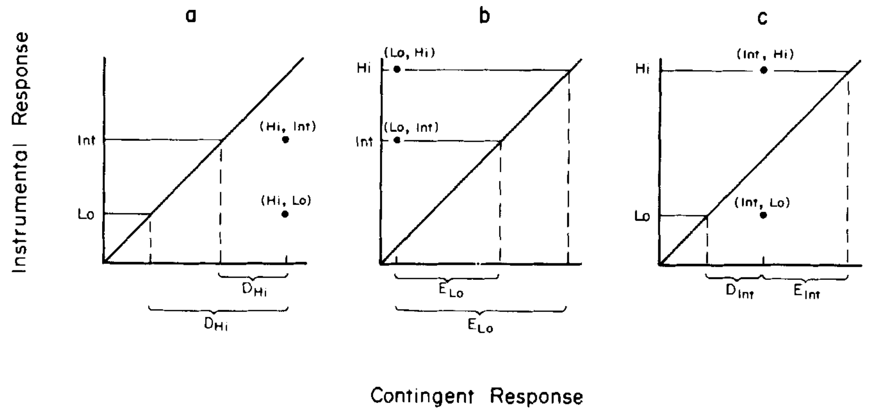


Fig. 3. Disequilibrium conditions that produce a reinforcement hierarchy for three responses. (a) A high-probability contingent response results in deficit conditions with the remaining responses. (b) A low-probability contingent response results in excess conditions with the remaining responses. (c) An intermediate-probability contingent response results in an excess condition with the higher probability instrumental response and a deficit condition with the lower probability instrumental response.

response, a one-to-one schedule produces a deficit condition for running. In Fig. 4b, where probability of drink is larger, the same schedule produces a deficit condition for drinking. Thus each response, in turn, should be increased.

7. The reinforcement relation can be reversed within a session by imposing a schedule before and after reversal of the relative baseline probabilities.

Figure 4 also illustrates this prediction. If Fig. 4a represents the first 5 min of a session, and Fig. 4b represents the last 5 min of a session, then during the last 5 min, contingent drinking should increase running, and during the first 5 min, contingent running should increase drinking.

The next two predictions differ from the preceding in that they hypothesize a functional relation between two variables. Such predictions are important because they place more constraints on the outcomes that satisfy the prediction, and, thus, are more sensitive to disproof. The previous predictions required only the first two or three assumptions of the equilibrium approach; the following predictions require a form of the fourth assumption relating amount of disequilibrium to change in responding.

8. The amount of increase in instrumental responding will be directly related to the probability-differential between the baselines of the contingent and instrumental responses.

Figure 5b shows the basis for this prediction using different prob-

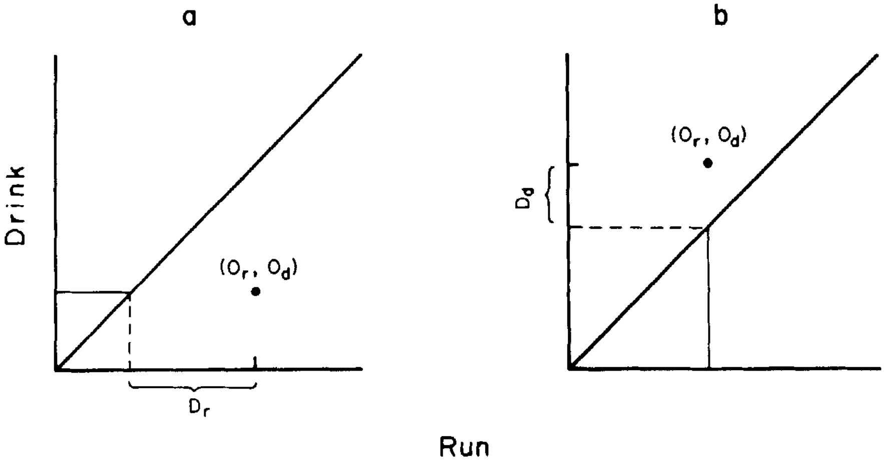


Fig. 4. Disequilibrium conditions that reverse the deficit condition by reversing the baseline probabilities of drink and run. (a) The combination of schedule and baseline that produces a deficit condition for run. (b) The combination of schedule and baseline that produces a deficit condition for drink.

abilities of the contingent response. As the baseline of the contingent response gets larger, the deficit condition for the contingent response gets larger. Thus, if the increase in instrumental responding is related to the size of the deficit condition, Premack's hypothesis can be accounted for.

9. The amount of increase in instrumental responding will be directly related to the baseline probability of the instrumental response (the base-increment hypothesis).

This prediction is confusing and contradictory from both the equilibrium and probability-differential viewpoints. The prediction of the equilibrium approach, shown in Fig. 5a, is clearly the reverse of Prediction 9. The higher the baseline of the instrumental response, the less the contingent response deficit, and, thus, the less increase in instrumental responding. Prediction 8 above, that the increase in instrumental responding should be a direct function of the probability-differential between the contingent and instrumental responses, also contradicts the base-increment hypothesis by predicting the same inverse relation between the baseline of the instrumental response and its increase under the schedule. An additional problem with the base-increment hypothesis is its counterintuitive prediction that the increase in instrumental responding will grow indefinitely large as the instrumental baseline approaches the contingent response baseline, until the probability-differential between them reaches zero. At this point the predicted increase in instrumental responding drops from infinity to zero.

The second point of confusion about this hypothesis lies in Premack's (1965) presentation. Though Premack (1965) clearly labels it the base-increment hypothesis, the data he cites as support refer most often to the absolute level of instrumental responding. His confusion of absolute and relative responding has been shared by others (Bauermeister, 1975; Burkhard, Rachlin, & Schrader, 1978). The confusion is important because the absolute level of instrumental responding may have little to do with schedule effects. For example, with a sufficiently large difference between the baselines of two instrumental responses, one could obtain no increase or even a decrease in the higher probability instrumental response under a schedule, and its absolute level of responding still could be higher than that of a low-probability instrumental response. This is not a very interesting prediction.

The third confusing aspect of this hypothesis is the data. Though the majority of the experiments support the existence of an inverse relation between increased responding and baseline of the instrumental response, some experiments clearly show the direct relation. Four studies showed an unambiguous inverse relation between the instrumental baseline and increment. Figure 6 shows the relation between the baselines of a series of manipulation responses for Chicko, Premack's (1963) prize Cebus monkey, and the increase in these responses under reward. Three of the responses led to one reward, and two responses led to another. Though the results are far from perfect, the trend is clearly for an inverse rather than a direct relation between the increase in the instrumental response

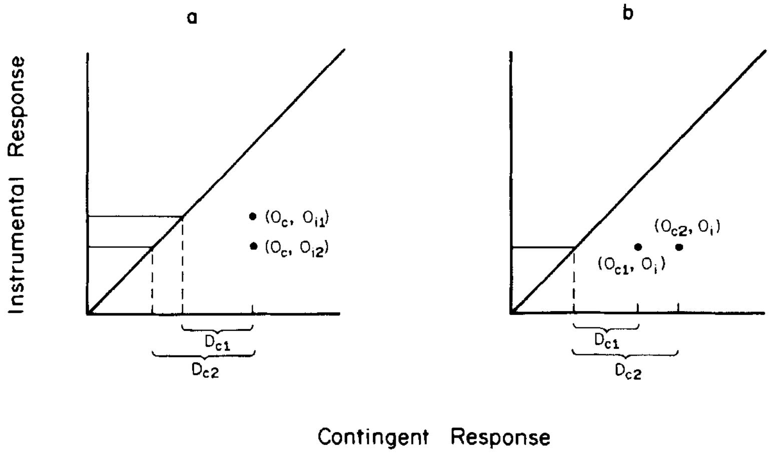


Fig. 5. Variation in the amount of contingent deficit as a function of baseline levels. (a) Different instrumental response baselines (O_i). (b) Different contingent response baselines (O_c).

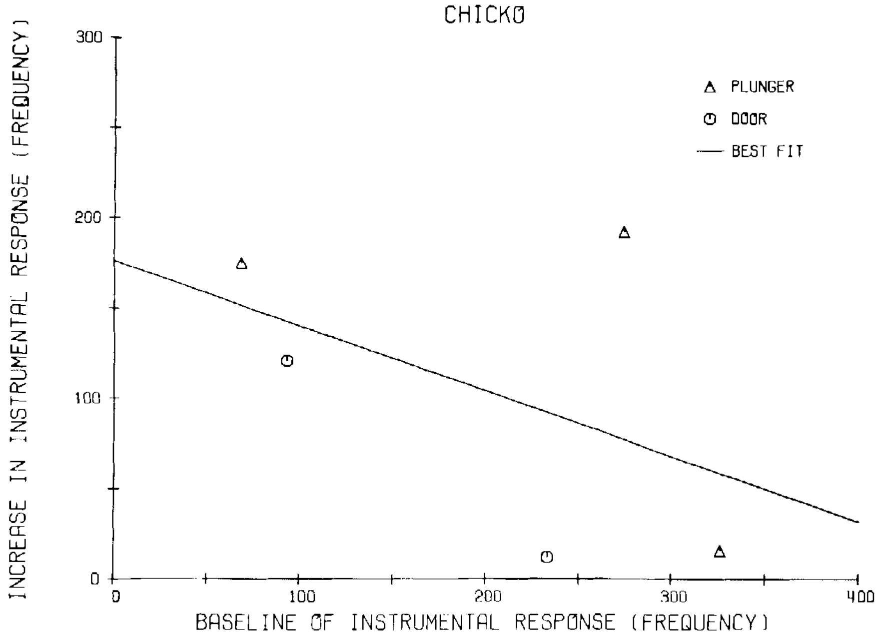


Fig. 6. Increase in instrumental manipulation responses in a Cebus monkey (Chicko) as a function of the baseline of the instrumental response. The line is a best fit regression line. The data are from Premack (1963).

and its baseline. A similar relation holds for two responses for Gimbel, another monkey in the same experiment.

Eisenberger *et al.* (1967), working with manipulation responses in humans, showed a marked inverse relation between baseline of instrumental lever pressing and the increase in instrumental responding for the opportunity to turn a wheel. Kjos (1977) tested the effect of the baselines of drinking, eating, and wood chewing on the increase in responding for access to a wheel in mice. He found a strong inverse relation between O_i/O_c (essentially O_i, since O_c was relatively constant) and the ratio of schedule to baseline performance ($r = -.61$). The correlation is much higher if the initial sessions of drinking are removed.

Wozny (1979) also found a marked inverse relation between baseline and increment in a study of water licking for wheel running in rats. Wozny (1979) manipulated the baseline of water licking by changing the water deprivation level, and subjects encountered all conditions. The results are displayed in a scatter plot in Fig. 7. It can be seen that for both the low and high baseline data, there is an inverse relation between individual baselines and increases in the instrumental response. Further, the average increment in instrumental responding was significantly greater for the low-baseline group, $t(13)=1.78$, $p<.05$.

Two studies appear to show a direct relation between instrumental baseline and increment, but on closer inspection there are alternative explanations of their data. Schaeffer (1965) manipulated the baseline of licking by using an empty tube, water, or 32 or 4% sucrose solution, and found that rats increased licking for access to a running wheel in direct proportion to the licking baseline. However, Schaeffer (1966) reported that the rats also overran and lost their limited access to the wheel in direct proportion to their baseline level of licking. Thus, Schaeffer's direct relation may have been due to the effect of the overrun in increasing the effective ratio schedule under which the animal was working. To test this possibility, I replicated Schaeffer's experiment using two conditions, one in which the animals were allowed to overrun the instrumental requirement, and one in which the scheduled ratio was assured by a reciprocal schedule that removed the instrumental response when the ratio was complete, and returned it only when the animal had run the appropriate number of wheel turns. The first group showed a direct relation between

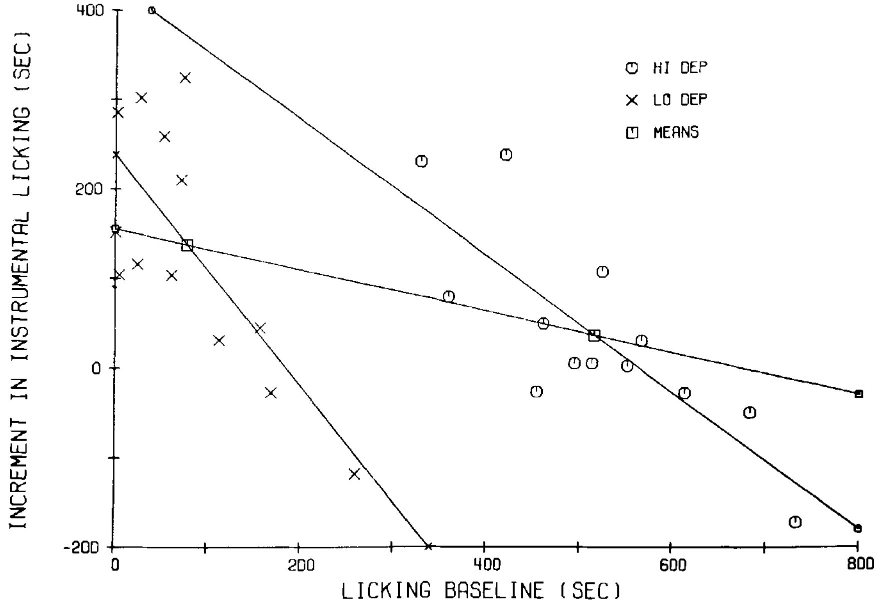


Fig. 7. Increase in duration of instrumental water licking for access to a running wheel in rats, as a function of individual differences in licking baselines and two levels of water deprivation (23.5 and 0 hr). Two of the lines are best fit regression lines. The line closer to horizontal connects the means of two deprivation conditions. The data are from Wozny (1979).

increment and baseline, and the second group showed an inverse relation between increment and baseline (reported in Timberlake & Allison, 1974).

Burkhard *et al.* (1978) reported evidence for a direct relation between instrumental increment and baseline in a study of children playing with toys. Children required to play with a toy of low baseline for access to a toy of high baseline showed a smaller increase than children required to play with a toy of intermediate baseline for access to the same high baseline reward (an average increase from 61 to 219 sec versus an average increase from 189 to 362 sec). Unfortunately these data are susceptible to several explanations. In both schedules there was a free toy available. For the low-probability instrumental response, the intermediate-baseline toy was freely available; for the intermediate-probability instrumental response, the low-baseline toy was freely available. Thus, the instrumental response of playing with the low-probability toy had more competition for expression than the instrumental response of playing with the intermediate-probability toy. This difference showed up in a markedly greater increase in playing with the free toy when the low-baseline response was instrumental. A second possible problem of interpretation was that the schedule requirements were not identical. The amount of play required for access to the high-probability toy varied from 60 to 90 sec.

The remaining studies showed both direct and inverse relations between instrumental baseline and increment, the last two in unambiguous fashion. Holstein and Hundt (1965) manipulated the baseline of barpressing within two rats by making brain stimulation or nothing contingent on barpressing in the baseline. In the contingency, when the barpressing package led to access to a sucrose solution, one animal showed a strong inverse relation between baseline and increment; the other animal showed a weaker direct relation.

Clear evidence for a direct relation between instrumental baseline and increase was provided by Bauermeister (1975) in a study of running for access to water in rats. Bauermeister (1975) manipulated the baseline of the instrumental response in two ways, by selecting animals that had high, medium, or low baselines of wheel running, and second, by imposing different degrees of torque on the wheel for three matched groups. In both experiments, Bauermeister imposed a series of fixed-ratio schedules of running for 10 sec access to the water.

Bauermeister reported a direct relation between baseline and increment for all schedules, but in several cases the effect was confounded with the overrunning effect noted by Schaeffer (1966). Figure 8 shows Bauermeister's results for the initial study redrawn to eliminate the percentage of instrumental responding that did not lead to the contingent response

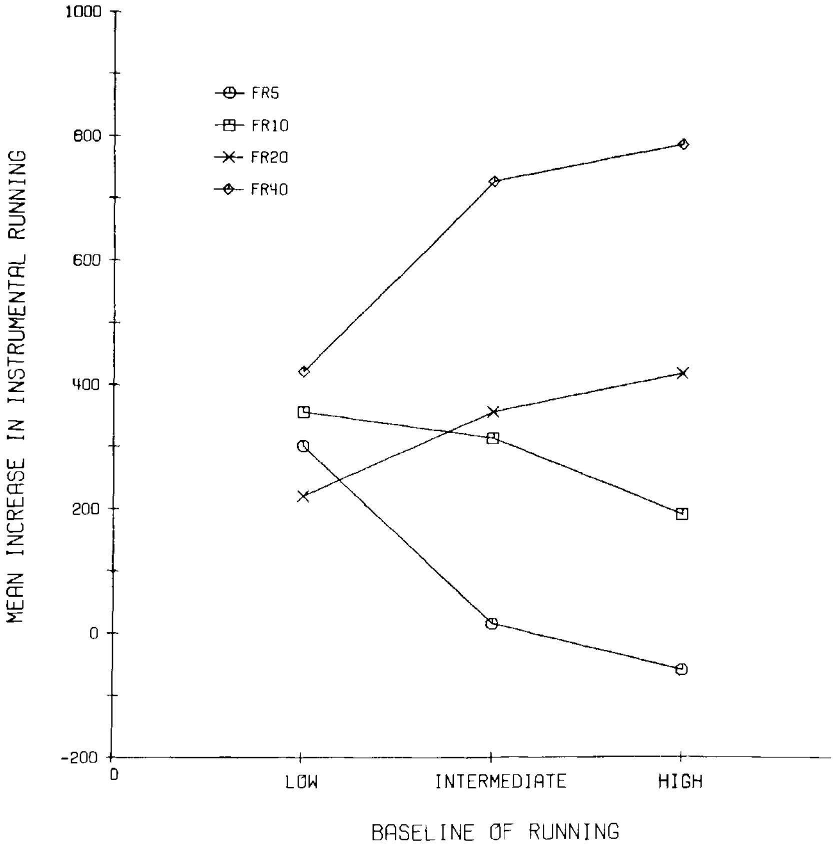


Fig. 8. Increase in frequency of half-wheel turns in rats for access to water, as a function of groups of animals with low, intermediate, or high baselines of wheel running. The data for four fixed-ratio schedules were adapted from Fig. 3 and Table 2 of Bauermeister (1975).

(shown in his Table 2). Though this procedure probably underestimates the amount of running that would occur if the subjects had been forced to drink, the difference is probably not great. The level of contingent licking under the schedule is very high under the low ratio schedules, probably approaching its baseline. Thus, little further increase in the instrumental response would be expected. Further, eliminating the amount of instrumental responding that did not lead to the contingent response does not consider that other instrumental responding led to a reduced amount of the contingent response due to partial overrun of the access period.

Redrawn in this fashion, Bauermeister's results are mixed. For low-ratio schedules, there was an inverse relation between increment and baseline. For the high-ratio schedules, there was a direct relation between increment and baseline. Figure 9 shows the results of the torque manipulation, also redrawn to eliminate the effects of overrunning the requirement. However, in this case the relation of increment to baseline is a direct one. Even if it were considered that the subjects used very little of the reward at low ratios, the evidence for a direct relation seems overwhelming.

Results similar to Bauermeister's first study were found by Wozny (1979) in an experiment in which different groups of rats ran for access to water under three different schedules, 24/24, 24/12, and 24/6 sec of running to sec of drinking. As indicated in the scatter plot in Fig. 10, the results for the first two schedules showed an inverse relation between baseline and increment, while the most extreme schedule showed the direct relation found by Bauermeister at higher ratio schedules and increased torque.

One possible explanation for these data is that at low ratio values (and low torque), the more important determinant of instrumental responding is the deviation of the contingent response from its baseline. Thus, the

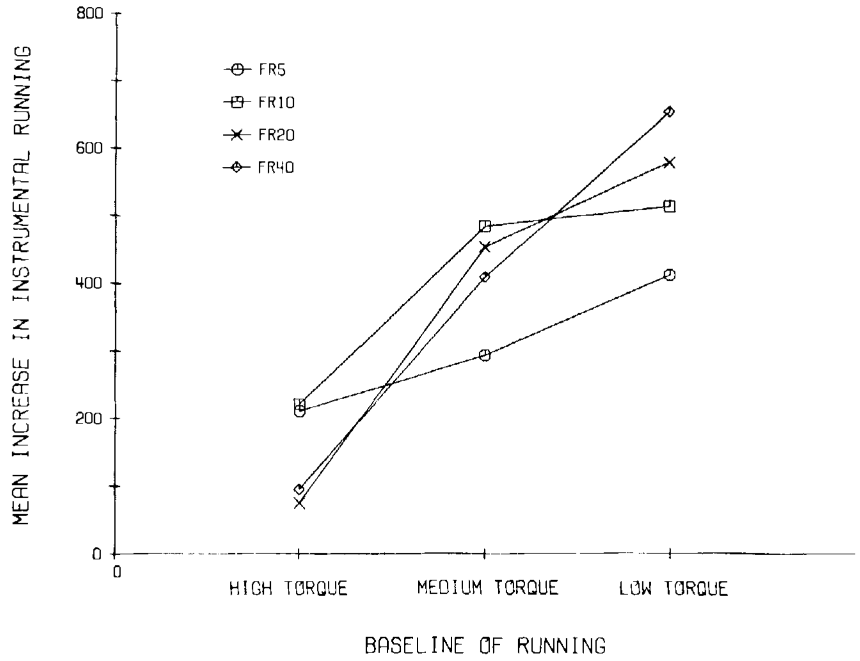


Fig. 9. Increase in frequency of half-wheel turns in rats for access to water, as a function of high, medium, or low torque on the running wheel. The data for four fixed-ratio schedules were adopted from Fig. 4 and Table 2 of Bauermeister (1975).

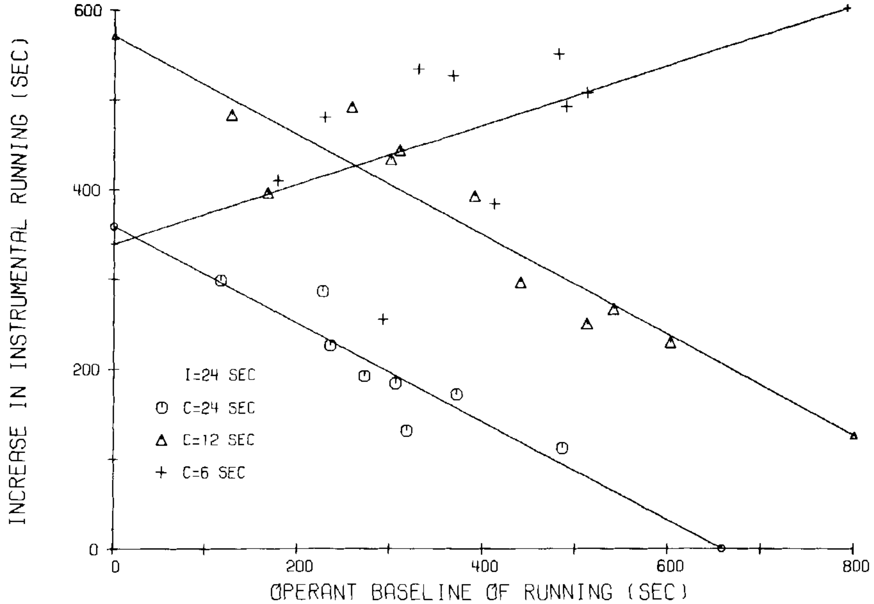


Fig. 10. Increase in seconds of instrumental wheel running for water in rats as a function of scheduled ratio and individual differences in baseline. The lines are best fit regression lines. The data are from Wozny (1979).

lower the instrumental baseline, the larger the deficit condition, and the greater the increase in the instrumental response. At high ratios (and/or high response costs), the deviation of the instrumental response from its baseline may be weighted more heavily. High-cost, low-probability responses, such as wheel running with high torque, may have greater resistance to increase than higher probability responses. Further, at large ratios high-probability instrumental responses are not required to increase as much for a given level of reduction in the contingent response. Under these circumstances, the relation between instrumental baseline and increment may be direct. A last factor of importance may be the experience of the organism. The inverse relation appeared more often in studies in which subjects encountered all conditions.

In summary, it appears that the equilibrium approach can account for all of the predictions of the probability-differential hypothesis, except the direct relation between the instrumental baseline and increment. In this case, the equilibrium approach seems to be accurate for low work requirements and within-subject manipulations of the baseline and schedules. The results for higher work requirements and larger deficit conditions appear to constitute an exception for a disequilibrium approach that focuses on the deviation from the contingent baseline, but may be assimilated with an approach that considers more carefully the importance of the deviation from the instrumental baseline.

B. RESPONSE CHANGE PREDICTIONS SPECIFIC TO AN EQUILIBRIUM APPROACH

As noted above, Premack's predictions are a restricted set of the predictions of an equilibrium approach. Many of these latter predictions contradict the predictions of the probability-differential hypothesis, while other predictions are unrelated to the Premack approach. The predictions in this section center on Assumptions 1-3, that a disequilibrium condition between baseline and schedule will produce instrumental behavior that tends to reduce this disequilibrium. Essentially most of these predictions follow readily fron the response deprivation approach outlined by Timberlake and Allison (1974) and its application to punishment schedules suggested by Heth and Warren (1978).

1. *Basic Equilibrium Predictions*

In the previous section I showed that, under certain conditions, the equilibrium approach is able to predict that a high-probability contingent response will increase a low-probability instrumental response, and the reverse contingency will produce no effect or punishment of the high-probability response. In this section I will show the opposite predictions also based on the equilibrium approach. Under appropriate conditions, a low-probability contingent response will increase a high-probability instrumental response, and a high-probability contingent response will not increase a low-probability instrumental response. The complete contrast of the predictions of the probability-differential and equilibrium approaches can be seen in Table II. For our purposes here, the contrasting predictions occur in the upper right and lower left squares.

Perhaps the most counterintuitive of the predictions that support only equilibrium theory is that a contingent response of lower probability can increase an instrumental response of higher probability. Figure 11a shows that as long as the scheduled ratio exceeds the baseline ratio, the nature of the baseline ratio is irrelevant in determining the presence of a deficit condition.

Though the evidence for this proposition has been slow to accumulate, it is now considerable. Eisenberger *et al.* (1967) first showed the effect in experiments on manipulation responses in humans (see Table III). Subjects in one experiment increased the higher probability response of turn-

TABLE II

CHANGES IN INSTRUMENTAL RESPONDING PREDICTED BY THE PROBABILITY-DIFFERENTIAL AND EQUILIBRIUM APPROACHES[a,b]

	Contingent deficit	Contingent equilibrium	Contingent excess
High-probability contingent response	+ +	+ 0	+ –
Equal-probability contingent response	0 +	0 0	0 –
Low-probability contingent response	– +	– 0	– –

[a] The upper left corner of each square contains the prediction of the Probability-Differential approach. The lower right corner contains the prediction of the Equilibrium approach.

[b] The plus means increase, 0 means no change, and minus means decrease if the schedule requires the subject to enter the negative state (perform the low probability or excess response) to obtain access to the other response. Without this requirement, the minus may refer only to no change.

ing a wheel for access to a lever pressing response, while subjects in a second experiment increased the higher probability knob manipulation response for access to a lever-pressing response. A control experiment showed that simple reduction in the lever-pressing response was insufficient to increase the instrumental response.

These results with manipulation responses in humans have been replicated by Klajner (1975), using lever pressing for access to a lower probability contingent response of pushing a plunger, and by Bernstein and Hinkle (1977), using the response of wheel turning for access to a low-probability contingent response of lever pressing. In both cases the higher probability instrumental response increased under the schedule.

Allison and Timberlake (1974) in separate experiments showed that rats would increase the higher probability response of licking a .4% saccharin solution for access to a lower probability response of licking either .3 or .1% saccharin. Measurement of a single baseline of licking the .4% solution indicated that the increase was not due to reduction in access to the contingent response.

Klajner (1975) showed that a higher probability response of water drinking in rats could be increased by contingent access to running in a wheel. Allison *et al.* (1979) found that a higher probability wheel running response was increased by access to a lower probability response of drinking .1% saccharin solution.

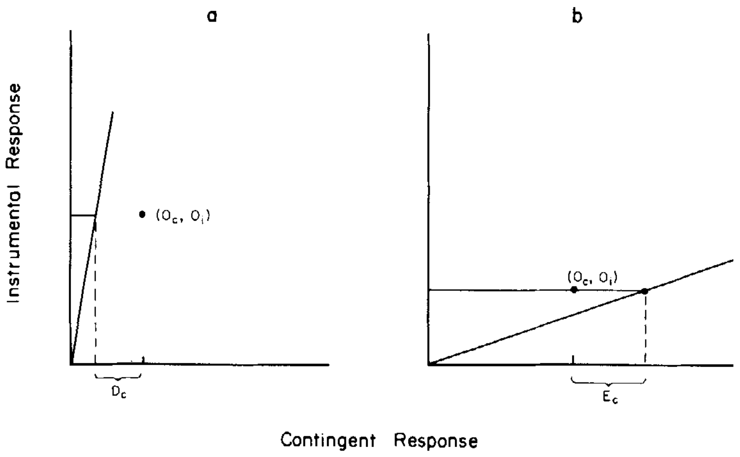


Fig. 11. Predictions from disequilibrium conditions that contradict the predictions of the probability-differential hypothesis: (a) A combination of schedule and baseline values that produces a deficit condition for a lower probability contingent response. (b) A combination of baseline and schedule values that produces an excess condition for a higher probability contingent response.

Dunham (1977) reported data from Marmaroff (1971) that showed that rats increased a more probable licking response for access to a less probable wheel-running response. In a subsequent experiment, Dunham reported the case of one animal that increased a more probable running response for access to a less probable licking response. In addition, Dunham (1977) reported that rats showing no preference between running and drinking readily increased either response depending on which response was relatively deprived by the schedule.

Bernstein and Dearborn (1978) investigated the effects of five different schedules relating low probability contingent responses and high proba-

TABLE III

AVERAGE FREQUENCY OF HIGH-PROBABILITY RESPONSE DURING BASELINES, AND WHEN INSTRUMENTAL IN GAINING ACCESS TO THE LOWER PROBABILITY CONTINGENT RESPONSE OF LEVER PRESSING[a]

Instrumental response	Paired baseline	Contingency	Paired baseline
One-quarter wheel turns (Exp. II, $n = 21$)	323	695	—
One-quarter wheel turns (Exp. III, $n = 8$)	226	106[b]	—
Knob manipulations (Exp. IV, $n = 12$)	62.8	181.8	126.0[c]

[a] Assembled from Eisenberger, Karpman, and Trattner (1967).

[b] Single baseline control for Experiment II.

[c] $n = 10$.

bility instrumental responses in humans. The results are displayed in Table IV. Though the size of the effect varies considerably, subjects showed consistent increases in the higher probability instrumental response. Further, as indicated in the bottom of Table IV, the size of this increase could not be attributed to either the reduction in contingent responding or its periodic access under the schedule.

Konarski (1979) and Konarski, Johnson, Crowell, and Whitman (1980) tested the effects of a lower probability response of working math problems contingent on the higher probability response of reading or coloring in normal or retarded grade school children. If the schedule satisfied the deficit condition, the children increased their amount of coloring behavior. Under the conditions of no contingent deficit, or a matched control, the children showed no similar increase in coloring or reading. Taken together these results strongly support the equilibrium position and contradict the probability-differential hypothesis.

The second equilibrium prediction that contradicts the probability-differential approach is that a higher probability contingent response does not guarantee an increase in instrumental responding unless the schedule produces a deficit condition for the contingent response. Figure 11b shows that even with a high-probability contingent response, a ratio line can be selected that produces a condition of excess for the contingent response. In the only separate studies of this hypothesis, Konarski (1979)

TABLE IV

INSTRUMENTAL RESPONDING IN HUMANS UNDER DEFICIT AND EQUILIBRIUM SCHEDULES IN WHICH THE BASELINE OF THE CONTINGENT RESPONSE WAS LOWER THAN THE BASELINE OF THE INSTRUMENTAL RESPONSE[a]

Schedule condition	Probability in baseline	Probability under schedule	Percentage change
Deficit			
S_1: Knit for yoga	.11	.56	409
S_1: Knit/sew for read	.57	.64	12
S_2: Read for instrument	.24	.28	17
S_2: Instrument for exercise	.11	.16	45
S_3: Art for sew	.14	.22	57
Equilibrium			
S_1: Sew for yoga	.47	.36	−23
S_2: Read for instrument	.27	.31	15
S_3: Art for sew	.24	.23	−4

[a] Assembled from Bernstein and Dearborn (1978).

and Konarski *et al.* (1980), using several school children who preferred solving math problems to reading, found that the children increased reading for access to solving math problems only if the schedule produced a deficit condition for contingent math. Increased math behavior was not shown when the deficit condition was absent. Data reported in the next section also support this prediction.

2. *Reversibility of Reinforcement by Schedule Changes*

Perhaps the most striking single prediction of the equilibrium approach is the reversal of the reinforcement relation between two responses by changing only the values of the schedule relating them. The baseline values remain the same. Figure 12 shows the reversibility prediction of the equilibrium approach using the responses of wheel running and eating (Timberlake & Wozny, 1979). It can be seen that the top-most schedule produces a deficit condition for eating, and, thus, should produce an increase in instrumental wheel running. The lower schedule produces a deficit condition for running, and, thus, should produce an increase in instrumental eating. If reciprocal schedules are used to relate the two responses, a test of reversibility of reinforcement by schedule changes simultaneously tests the predictions in the four corner boxes in Table II. Essentially, an increase in instrumental responding is predicted if, and only if, a contingent deficit condition is produced by the schedule.

Several experiments have shown reversibility of reinforcement by schedule changes across different experiments. Eisenberger *et al.* (1967) showed that pressing a lever could serve either as the instrumental re-

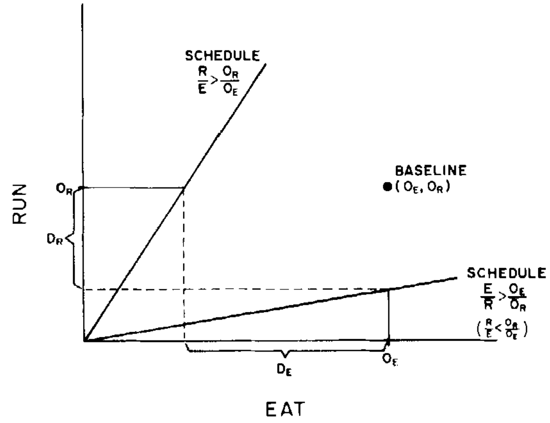


Fig. 12. Disequilibrium conditions that reverse whether eating or running is under a deficit condition by changing the scheduled ratio.

sponse for contingent wheel turning, or as the contingent response for instrumental wheel turning, depending only on schedule changes. Konarski *et al.* (1980) showed that working math problems would either increase reading, or itself be increased by the opportunity to read, again depending only on changes in the schedule.

Other experiments have tested reversibility in a single experiment. Timberlake and Allison (1974) and Mazur (1975) showed reversibility of reinforcement for saccharin drinking and wheel running in rats using the same subjects, and altering only the nature of the reciprocal schedule relating the two responses. More recently, Timberlake and Wozny (1979) and Wozny (1979) have shown reversibility of reinforcement between eating and wheel running and between water drinking and wheel running in rats (see Fig. 13). The latter experimenter employed both a within and between-subjects design. Heth and Warren (1978), using humans, showed reversibility of reinforcement by schedule changes between pressing buttons for auditory and visual stimulation. Last, Podsakoff (1980), using humans in a complex factorial design, showed clear reversibility by schedule changes between trigger pulling and pursuit rotor performance.

Two of these studies did not include a test for the noncontingent effects of schedules, a factor that can be of some importance when using highly probable or substitutable instrumental responses (Timberlake, 1979). However, Konarski *et al.* (1979), Timberlake and Wozny (1979), and Wozny (1979) used a combination of matched and massed baselines to determine noncontingent effects of the schedules. Eisenberger *et al.* (1967), Heth and Warren (1978), and Podsakoff (1980) also considered increases due to noncontingent factors. Last, there is one reported failure to obtain reversibility between wheel running and water licking in mice under 24-hr running conditions (Ebbesen, Allen, & Kjos, 1977). The reasons for this result are not yet clear.

3. *Maintenance of Baseline*

An important prediction of a simple equilibrium approach is that the subject, given the opportunity, will maintain its baseline response levels, that is, a condition of no deficit or excess. The basis for this prediction can be seen by imagining a schedule line drawn through the baseline point in any of the above figures. No condition of excess or deficit occurs, and, thus, no change in responding is predicted. This prediction should be contrasted with the predictions of the probability-differential approach in Table II.

Evidence for this prediction is moderate, though generally favorable to equilibrium theory. Premack (1965) and Jacobson and Premack (1970)

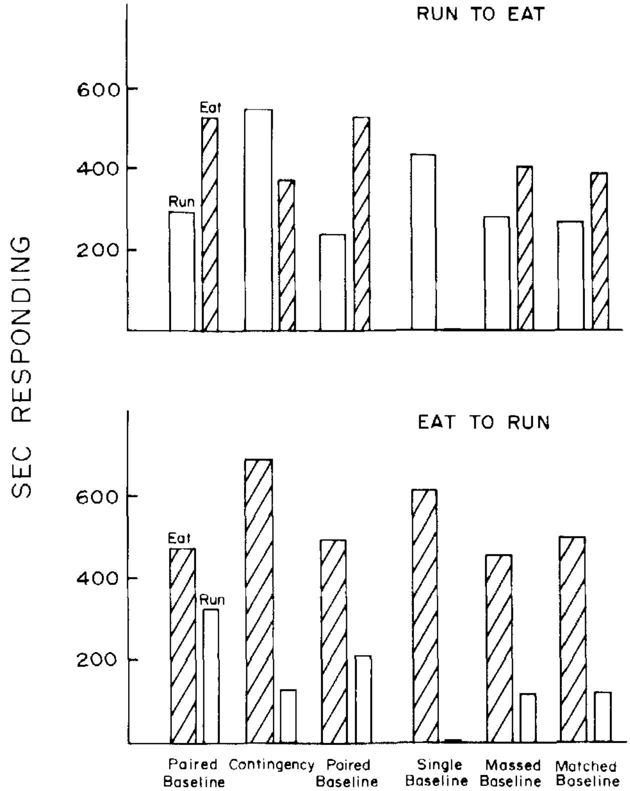


Fig. 13. Seconds of wheel running and eating under paired baselines and contingencies, and single, massed, and matched baselines. The response under the deficit condition is listed second. Instrumental responding under the contingency significantly exceeded all baselines. The data are from Timberlake and Wozny (1979).

reported obtaining the phenomenon, using wheel running and drinking in rats, but it is difficult to evaluate their data. Timberlake (1974) and Wozny (1979), using wheel running and saccharin licking, and wheel running and water licking, respectively, found the effect for small numbers of subjects. Figure 14 shows the results of their experiments on the same plot, as a function of percentage increase or decrease in baseline responding under the baseline. Though the variability is higher for the undeprived animals, both sets of scores bracket the baseline.

Peden and Timberlake (1976) and Timberlake and Peden (1977), in studying the keypecking of pigeons for food, performed an approximate test of this hypothesis. They required pigeons to complete a fixed-ratio 10 or a variable-interval 30-sec schedule for access to their baseline level of grain intake. As indicated in Table V, under these conditions the birds approximated their baseline levels of keypecking and eating. Inspection of Premack's (1963) data on manipulation responses in Cebus monkeys also indicates that when the schedule approximated the baseline ratio, the monkeys maintained their baselines (though there were clear exceptions).

Positive evidence for baseline maintenance in humans comes from the work of Bernstein and Dearborn (1978; bottom of Table IV), Konarski (1979), and Konarski *et al.* (1980). These researchers mixed schedules that produced a deficit condition with equilibrium schedules that produced no deficit. In all equilibrium schedules there was no systematic change in instrumental responding. The last positive evidence in humans comes from Podsakoff (1980) who studied the effect of equilibrium schedules on three separate groups that differed in the size but not the

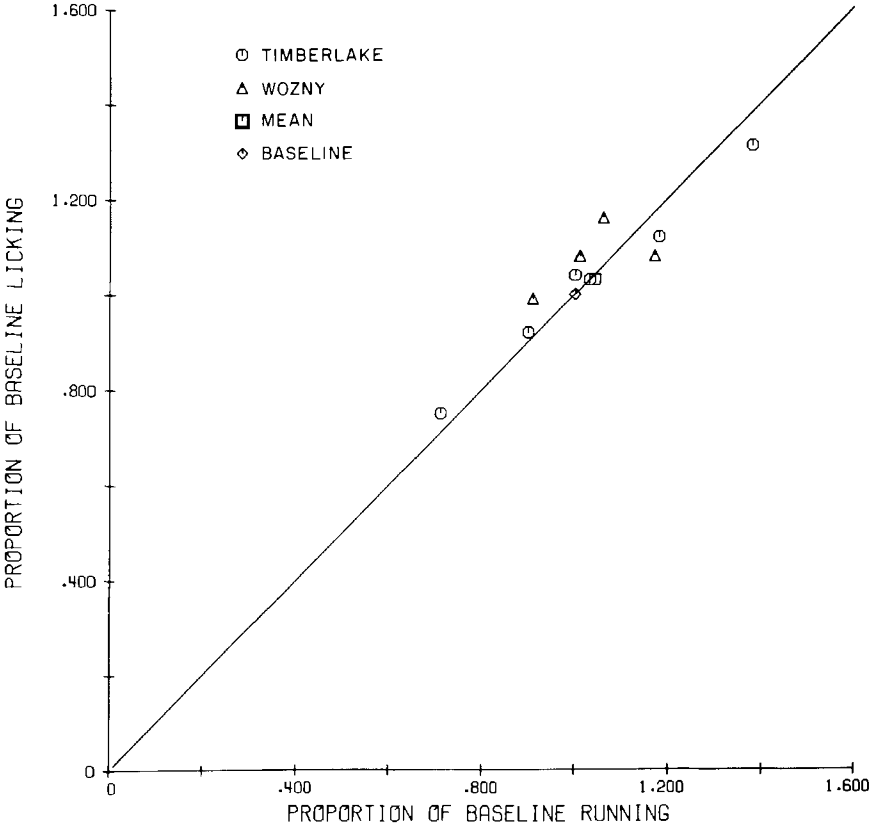


Fig. 14. Proportion of baseline licking and wheel running in rats under schedules for which the baseline and scheduled ratios were equal. The data are from Timberlake (1974) and Wozny (1979).

TABLE V

KEYPECKING AND EATING IN PIGEONS UNDER BASELINE AND FIXED-RATIO 10 AND VARIABLE-INTERVAL 30-SEC SCHEDULE REQUIREMENTS FOR ACCESS TO UNLIMITED HOPPER TIMES[a]

	Baseline		Contingency	
Subject	Minutes eating	Keypecks	Minutes eating	Keypecks
FR 6940	3.8	15.7	3.8	10.8
FR 3601	8.9	14.9	8.0	12.0
VI 7377	6.8	0.0	6.3	3.3
VI 6667	14.9	13.7	15.8	5.7
VI 6777	14.6	3.9	13.3	1.3

[a] Assembled from Timberlake and Peden (1977).

ratio of the schedule terms. None of the groups showed a significant change in trigger pulling or pursuit rotor performance.

Two exceptions to maintaining baseline have been reported. Bernstein and Hinkle (1977), in studying manipulation responses in humans, found two groups that appeared to violate this prediction. However, their data are readily susceptible to alternative interpretation. One of the groups showed a decrement in the contingent response accompanying the increment in the instrumental response, thereby showing that the effective schedule was not an equilibrium schedule. The second group increased contingent responding from zero in the first baseline to very high in the following baseline, thereby indicating that the instrumental baseline was not stable, perhaps due to changing demand characteristics in the situation.

The second exception was reported by Mazur (1975) in his data on wheel running and licking in rats. Figure 15 shows that the subjects all performed less of both responses under the schedule than they did under the baseline condition. Allison *et al.* (1979) suggested that Mazur's apparatus may have affected the outcome of his experiments. Since the instrumental and contingent responses were separated by a considerable distance (more than 35 cm), the schedule required considerable movement, back and forth, that was not required in baseline. Allison *et al.* (1979) showed that requiring a long trip between two responses produced a lower level of the instrumental response than did a shorter trip. Similar results have been reported by Roper (1973,1975). This effect may occur because the instrumental requirement in the schedule underestimates the effective requirement, or because associative aspects of the situation have been changed.

Though the results generally support the equilibrium prediction, they also suggest the need for more careful research. Effects such as those reported by Schaeffer (1966), in which rats licking sucrose overran access to a contingent running response, suggest strongly that the local organization of behavior (e.g., the burst length of licking sucrose) can affect the outcome of a disequilibrium condition. It would be surprising if such results could not be found in schedules producing zero molar disequilibrium.

4. *Punishment*

The traditional view of punishment is that it is the forcible imposition of a noxious state of affairs contingent on the performance of a particular

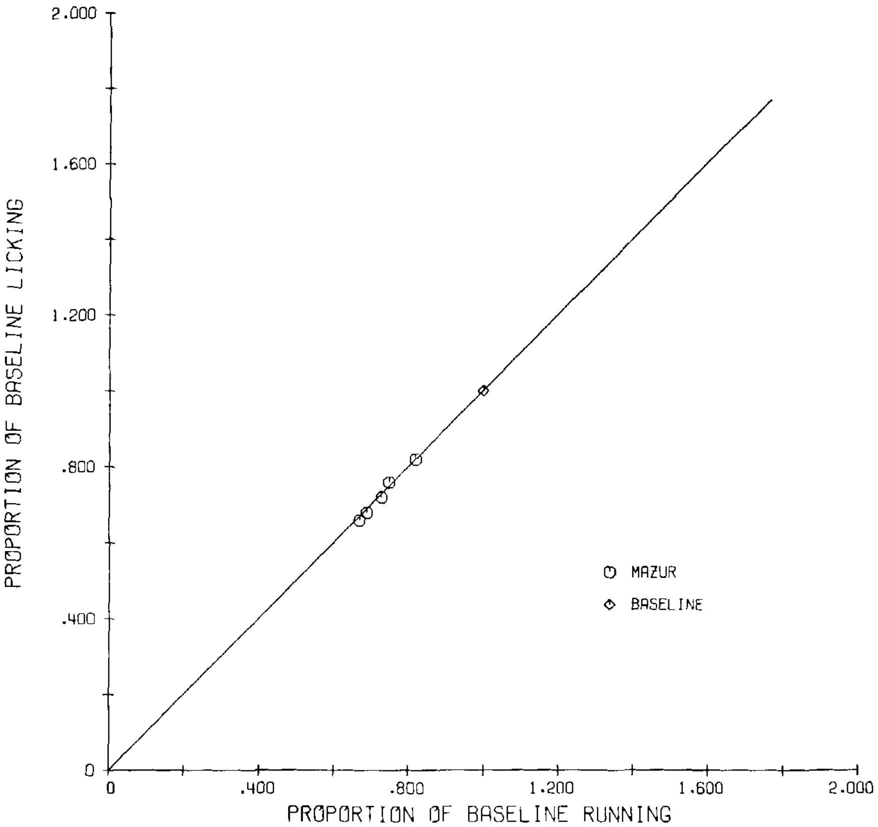


Fig. 15. Proportion of baseline licking and wheel running in rats under schedules for which the baseline and scheduled ratios were equal. The data are from Mazur (1975).

behavior. In the probability-differential view, punishment is produced by a low-probability contingent response that is forced on the subject. In the equilibrium view, at least in theory, there are no intrinsically noxious states, and the probability of the contingent response is irrelevant (see the right column of Table II). A noxious state is simply a state that the subject is forced into at a rate greater than its baseline, thereby producing a state of forced excess (see Fig. 1b; see also Premack, 1971, p. 148).[7]

At a procedural level there are two ways to produce a condition of forced excess. The traditional technique is to physically impose a state on the subject, such as shock or loud noise. A second way is to impose a schedule that allows the subject to choose whether to exceed its baseline of one response to gain access to its baseline of the other. In the latter case the subject can choose not to increase the excess response, but unless it does so, it will not regain access to the deficit response. From the viewpoint of equilibrium processes alone, whether the experimenter or the schedule "forces" the subject, the prediction is the same. In actuality, there may be differences in the subject's reaction to these conditions because of associative variables, but the prediction of decreased instrumental responding remains unchanged.

This rather broad view of punishment has important implications for schedules that are traditionally thought to deal only with appetitive effects. In any disequilibrium schedule in which the subject must perform one response to gain access to the other, there exists both a deficit condition and an excess condition (see Fig. 1a). Thus, from the equilibrium viewpoint, a subject expected to increase its instrumental response to reduce the deficit condition for the contingent response will also be expected to decrease its contingent response to reduce the excess condition for the instrumental response. The only exception to this prediction occurs if the schedule does not force the subject to perform the response in relative excess to obtain access to the response in relative deficit. The subject may then ignore some of the opportunity to increase the response in relative excess (e.g., a subject wheel runs for excessive access to food, but need not use the access to run again).

It follows from the above analysis that data contradicting the probability-differential predictions for punishment (shown in the bottom of Table II) have been collected already in the form of data contradicting the probability-differential predictions for reinforcement. In a reciprocal schedule, if a contingent response of low probability increased an instru-

---

[7] In practice there may be a problem in determining the relative instigation of responses having a zero baseline. Other procedures may be necessary to distinguish their relative attractiveness (e.g., Timberlake & Allison, 1974).

mental response of high probability, then the instrumental response of high probability probably decreased the contingent response of low probability. Likewise, if a contingent response of high probability decreased an instrumental response of low probability, then the instrumental response of low probability probably increased, rather than decreased, the contingent response of high probability.

The data reviewed in the previous section support these predictions quite clearly. The general prediction that a schedule producing a disequilibrium will both increase and decrease responding also has been supported in wide-ranging reviews of the literature by Allison (1976) and Allison *et al.* (1979). As we will see, there are data that violate this pattern of results, but not in a way that supports any present alternative account.

5. *Location*

As can be seen from the above arguments, the equilibrium approach has rather strong implications about where responding under a schedule ought to fall. The limits on responding are indicated in Fig. 16 by a set of lines parallel to each axis and passing through the baseline point. Under a deficit condition for the contingent response, responding should fall in Quadrant 1. Under an excess condition for the contingent response, responding should fall in Quadrant 4. Despite the fact that the schedule lines go through Quadrants 2 and 3, responding should not fall in these quadrants because these points involve increased deviations from baseline for both responses.

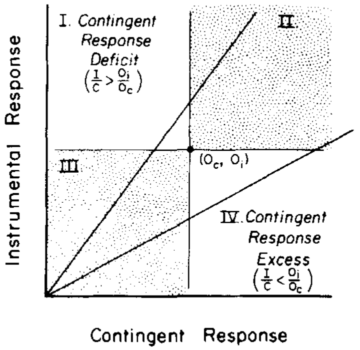


Fig. 16. Theoretical constraints on the location of responding under schedules that produce conditions of deficit or excess. No responding should fall in the shaded areas. Under schedules that produce a condition of deficit for the contingent response, responding should fall in Quadrant I. Under schedules that produce a condition of excess in the contingent response, responding should fall in Quadrant IV.

In general, the data reviewed above support these predictions in the case of positive ratio schedules. However, there appears to be a systematic exception to this finding. Subjects working under schedules that produce a relatively small deficit condition often increase both instrumental and contingent responses so that responding falls in Quadrant 2 of Fig. 16. Figure 17 shows the behavior of two pigeons keypecking for a percentage of their daily food intake under a series of ratio schedules (Timberlake & Peden, 1977). At ratios slightly above the baseline ratio, both pigeons exceeded the amount of keypecking and feeding in baseline. One pigeon continued to respond to Quadrant 2 for all ratio schedules run.

Similar effects are scattered throughout the literature. Hogan and Roper

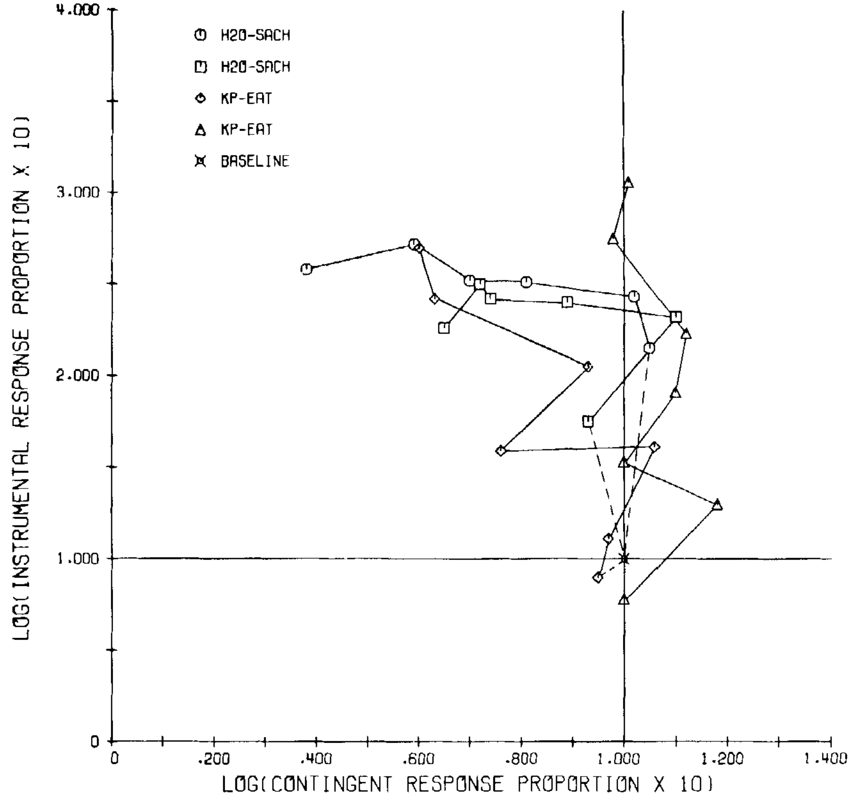


Fig. 17. Keypecking and eating in two pigeons and water and saccharin licking in two rats as proportions of baseline responding. The different points for each animal represent responding on a series of ratio schedules that produced increasing deficits for the contingent response. The pigeon data are from Timberlake and Peden (1977). The rat data are based on two of six animals James Wendling and I ran.

(1978) in their extensive review of regulation in learned performance showed several examples of Quadrant 2 responding on low ratio schedules. The effect is also seen in the data of Marwine and Collier (1979), using rats barpressing for water, and Kanarek (1976), using rats barpressing for food. The basic effect is not confined to the use of low-probability instrumental responses. Data from Wozny (1979) show the effect in rats drinking water for access to a wheel (see Fig. 18). Nor is the effect confined to the use of deprived animals and food and water reward. Klajner (1975) found the effect in the case of humans pressing a lever for access to a plunger, and I found it in the case of undeprived rats drinking water for access to a saccharin solution (see Fig. 17 for the data of two

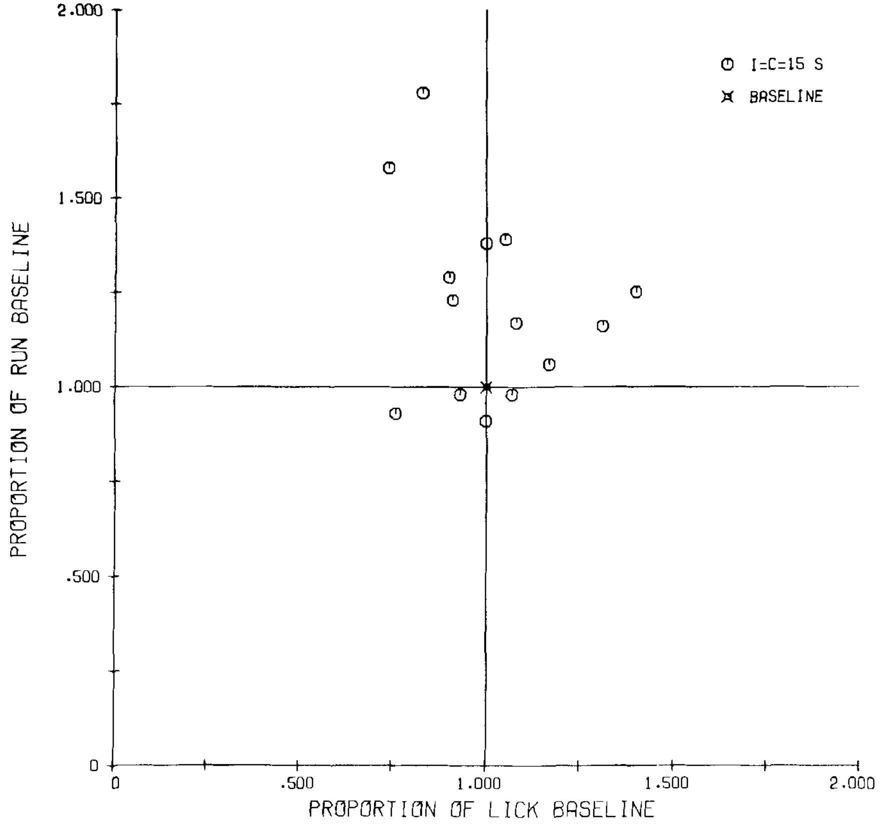


Fig. 18. Proportion of baseline wheel running and licking in rats under a schedule that required 15 sec of running for 15 sec of drinking. This schedule produced a slight condition of deficit in drinking for most of the rats. The data are from Wozny (1979).

subjects). The effect also occurs in interval schedules in pigeons keypecking for food (Timberlake & Peden, 1977).

The reasons for this exception to the predictions of equilibrium theory are not clear. In the case of intake of food and water, it could be argued that the animal was taking in more food or water to compensate for its increased work load under the schedule. But this does not explain why the subjects did not continue this behavior at higher ratios, or why commodities as unimportant as plunger pressing or saccharin licking in undeprived animals produced the effect as well. It may be that periodic access to the contingent response alters total instigation in the situation (Falk, 1971), or that the subject resists schedule disruptions of baseline responding (Bernstein & Dearborn, 1978; Kavanau, 1969). The effect does not always occur, but its frequency appears to warrant further study.

C. FUNCTIONAL PREDICTIONS OF EQUILIBRIUM THEORY

The predictions reviewed above require only that there exist tendencies to reduce the disequilibrium condition imposed by the schedule (and a salient and potentially effective instrumental response). The predictions in this section require that there be a functional relation between measures of disequilibrium and changes in instrumental responding. The working hypothesis, that the increase in instrumental responding will be a direct function of the amount of initial disequilibrium imposed by the schedule, has considerable support (Allison, 1976; Hogan & Roper, 1978; Podsakoff, 1980; Terhune, 1978; Timberlake 1977; Timberlake & Allison, 1974).

However, we have already seen that this hypothesis breaks down at large amounts of contingent deficit, and when the instrumental response is very costly. In this section, I will briefly review two other types of data that also show limitations of this assumption.

1. Equal Schedule Ratios, Unequal Size of Schedule Terms

A clear derivation from the disequilibrium hypothesis is that, for given baseline levels of responding, the amount of change in instrumental responding under a schedule will depend only on the ratio of the instrumental requirement to the contingent payoff, and not on the size of these schedule terms.[8]

---

[8] There will be limits on the size of schedule terms suitable for this prediction. An intuitive upper limit on the size of the instrumental requirement is that it not exceed the single baseline of the instrumental response; similarly, the size of the contingent pay-off should not exceed its paired baseline. Otherwise, the subject will not make good contact with the terms of the schedule.

Allison and Timberlake (1974) provided evidence supporting this view in a study of rats licking one solution of saccharin for access to another. Though the size of the terms in the ratio schedule varied by a factor of four, there was no significant difference in instrumental responding. Allison *et al.* (1979) provided similar evidence in the case of rats barpressing for access to water, as did Wasik (1969), using humans pressing levers to accumulate points on a variety of different ratio schedules. Podsakoff (1980), using humans and the responses of trigger pulling and performing on a pursuit rotor, also showed no effect of the size of the schedule terms at several different schedule ratios.

However, two studies that manipulated the absolute value of the schedule terms across a very wide range found systematic changes in instrumental responding. Kelsey and Allison (1976), in testing rats barpressing for food, found a marked curvilinear relation between the absolute value of the schedule terms and instrumental performance. Data from Timberlake and Peden (1977) also show a mixed relation between abso-

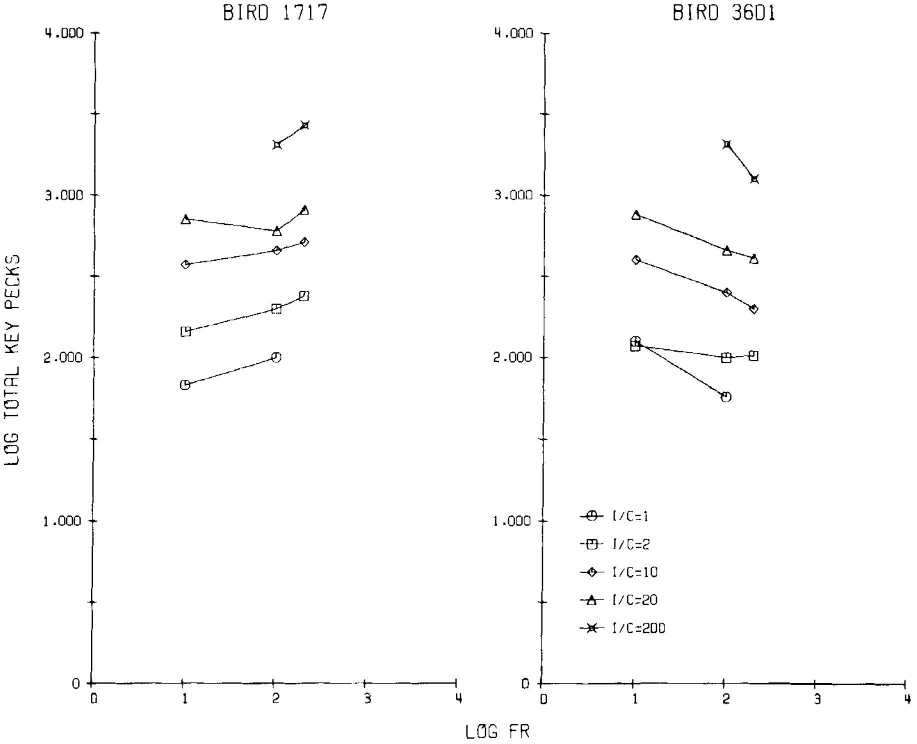


Fig. 19. Log keypecking for food in two pigeons as a function of the log of different requirements of keypecking. All points connected by a line represent equal ratios of the keypeck requirement to amount of food payoff. Data are from Timberlake and Peden (1977).

lute value of the schedule terms and amount of increase in instrumental responding (see Fig. 19). Using pigeons keypecking for food across a wide range of keypeck requirements and payoff values, they found one bird that systematically decreased responding with absolute size of the schedule terms, and another bird that systematically increased responding over the same variable.

2. *Schedule Changes*

The basis for the prediction of increased change in instrumental responding with increased difference between the schedule ratio and the baseline ratio can be seen in Fig. 20. As the scheduled ratio deviates from the baseline ratio, the amount of deficit or excess increases, and thus the change in instrumental responding is predicted to be greater. As indicated in the introduction to this section, the evidence for this prediction is very large, and it will not be reviewed here. Instead, I will consider data that suggest important limits on the accuracy of the prediction.

There is considerable evidence that as the schedule requirement increases, the amount and rate of instrumental responding first increases and

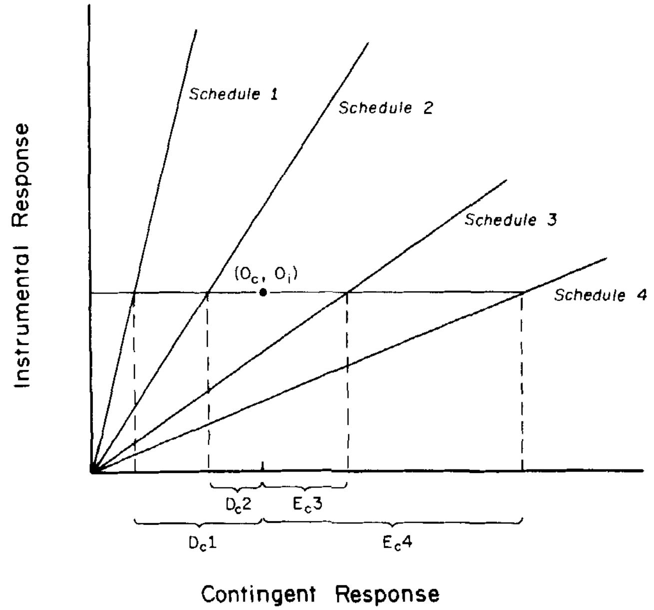


Fig. 20. Amount of deficit and excess conditions as a function of the difference between the scheduled ratio and the baseline ratio of the instrumental and contingent responses.

then decreases (Staddon, 1979; Timberlake, 1977). The increase in responding is congruent with the disequilibrium hypothesis; the decrease is not, unless the increasing effect of instrumental excess is considered. Timberlake and Peden (1977) showed evidence for a similar nonmonotonic relation between scheduled ratio and instrumental performance by manipulating the amount of access to grain while leaving the peck requirement constant. Figure 21 shows that at low ratios the relation between keypecking and payoff is inverse as predicted by the disequilibrium hypothesis. At high ratios and low rewards the relationship is a direct one. Figure 22 shows more data for the critical area where the shift in relation occurs (Peden & Timberlake, 1976).

3. Discussion

Though this section has been brief, the tentative conclusions should be clear. Considerable data support the disequilibrium hypothesis. Across a variety of manipulations, the change in instrumental responding is di-

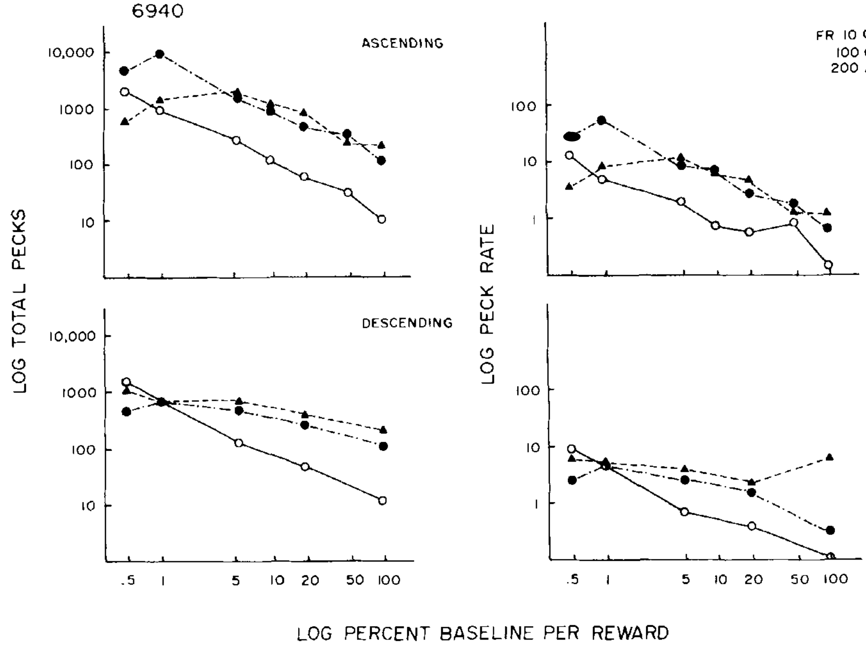


Fig. 21. Log total keypecks and log rate of keypecking for food in a pigeon as a function of the log percentage of daily ration given as payoff for completing each ratio. The three lines represent different ratio schedules. The data are a representative bird from Timberlake and Peden (1977).

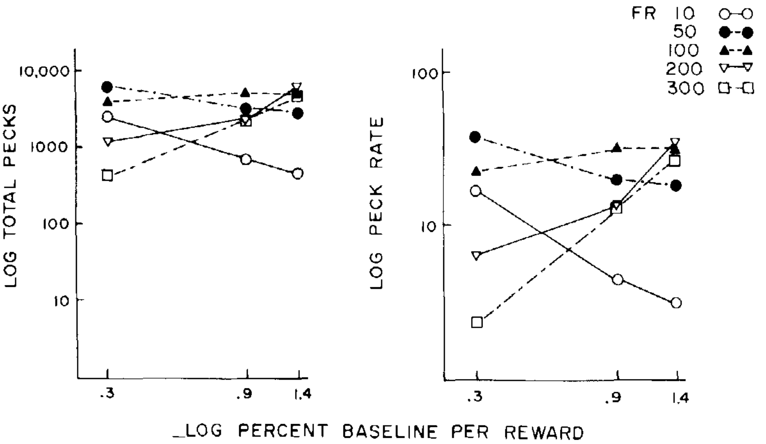


Fig. 22. Log total keypecks and rate of keypecking for food in a pigeon as a function of the log percentage of daily ration given as payoff for completing each ratio. The five lines represent different ratio schedules. For this bird, the size of the payoffs given was much smaller than for the bird in Fig. 21. The data are from Peden and Timberlake (1976).

rectly related to the size of the initial disequilibrium. However, the disequilibrium hypothesis in its present form appears inadequate to predict changes under some conditions. In a sense, the size of the initial disequilibrium represents the point at which the subject begins altering its behavior under the schedule. How the subject resolves the disequilbrium will depend on the nature and functional characteristics of the equilibrium processes underlying responding.

There are at least five classes of variable that are not dealt with adequately in the present equilibrium theory. Other theorists have considered some of these variables, but no general treatment yet accounts for the data. The first variable is the importance of the deficit in the contingent resonse. Lockard (1964) and Premack and Collier (1962) pointed out that responses differ considerably in how they are affected by deficits. Some tend to recover set points, while others over- and undershoot previous response levels. These characteristic differences should be relevant to behavior under a schedule. Even within a single response, it seems reasonable that the importance of a decrease in a response below its baseline will vary with the size of the decrease, probably in a nonlinear fashion (Timberlake & Wozny, 1979).

The second variable is the importance of the excess of the instrumental response. Some responses appear to cost more than others, a difference that is not necessarily reflected in baseline responding. Increasing the torque requirement on a running wheel (Bauermeister, 1975) or the force to press a bar (Allison *et al.*, 1979) produces greater resistance to instrumental increase. The importance of excess may also vary with the size of the deviation from baseline, and with any special relationship between the instrumental and contingent responses (Shettleworth, 1972; Allison *et al.*, 1979).

The third variable not adequately dealt with is the role of associative processes. Timberlake (1979, and in unreported data) has shown that the presence of a schedule increases the amount of instrumental responding for a given amount of deficit in the contingent response over what occurs if the response is merely reduced to a comparable amount. Other evidence suggests that this relationship changes as a function of the scheduled ratio, and perhaps with the size of the schedule terms.

The fourth variable affecting response changes is the presence of more complex equilibrium conditions, involving responses other than the instrumental and contingent responses. Bernstein and Ebbesen (1978), Burkhard and Rachlin (1978), and McIntosh (1974) showed that the increase in instrumental responding under a deficit condition depended on the other responses available to the subject. If responses similar to the contingent response were freely available, the effect of the deficit condition on instrumental responding was markedly reduced by substitution of the alternative response. If responses similar to the instrumental response were available, the increase in instrumental responding was sometimes enhanced. It may also be that responses outside the schedule session are relevant to responding under the schedule.

The last variable potentially affecting response change is the molecular structure of behavior. There may exist mutually excitatory or inhibitory effects among responses, either dependent on or independent of the temporal and sequential framework of the schedule. Further, the structural integrity of response characteristics, such as burst length or rate, may change the effectiveness of a schedule. It may also be that the temporal unit over which the animal integrates a disequilibrium condition influences overall response changes. In Section IV, I will briefly discuss two attempts to model the effects of molecular structure on responding.

IV. Molecular Determinants of Learned Performance

As attested by Dunham's (1977) article, the relation between changes in the molecular structure of responding and the overall molar changes in responding under a schedule is of considerable current interest and importance. Unfortunately, not much progress has been made in precisely determining the relation between molecular and molar responding.

A. MOMENTARY PROBABILITY

In 1971, Premack introduced the concept of momentary probability to refine and develop further his position on reinforcement. The concept of momentary probability has great intuitive appeal, but a history of multiple definitions and inadequate testing (Dunham, 1977; Timberlake & Wozny, 1979). The original and clearest use of the concept was to refer to local fluctuations in baseline probabilities. Premack (1971) and Bauermeister and Schaeffer (1974) showed that in the case of two responses that reversed in probability within a baseline session, the same schedule could produce reinforcement of either response, depending on when the schedule was imposed. Terhune (1978) showed that a baseline session could be divided more finely into six sections, and the relative increase in instrumental responding during each section could be predicted from the local baseline probability of the contingent response. From the equilibrium viewpoint, these findings demonstrate that the subject is able to respond to some local conditions of disequilibrium. If this supposition is true, it should be possible to reverse the direction of reinforcement findings by appropriate changes in the scheduled ratios (see Table II).

Premack also suggested that momentary probability might capture the difference between the attractiveness of behaviors that occur intensely at infrequent intervals, versus behaviors that occur less intensely, but more regularly. On consideration, this suggestion does not seem entirely reasonable unless one somehow measures the differences in intensity.

A third use of momentary probability is to designate absolute proabilities under a schedule. In a note to James Allison and me (Timberlake & Allison, 1974; Timberlake & Wozny, 1979), Premack suggested that the momentary probability of a response under a schedule referred to the amount of its baseline that was not yet expressed. In this view, the reinforcer at any particular time would be the response with the higher unexpressed baseline. This ingenious hypothesis combined the concepts of momentary probability (unexpressed baseline) and probability-differential to predict that a contingent response of lower baseline could reinforce an instrumental response of higher baseline, because at some point in the schedule session, the relative size of the unexpressed baselines would reverse (Timberlake & Allison, 1974).

Though intriguing, this view suffers from two major problems. First, it does not explain why no change in responding occurs when the scheduled ratio equals the baseline ratio, and there is a higher probability contingent response (Timberlake & Allison, 1974, this article). The second problem is that in schedules with a low-probability contingent response, the data do not support reasonable deductions relating changes in the pattern and amount of responding to the point of reversal in size of the unexpressed baseline (Timberlake & Wozny, 1979).

The final use of momentary probability was expressed by Premack in the last section of his 1971 paper. He claimed, "Operationally, reinforcement is produced by denying the subject the opportunity to occupy a state as long as it would choose to" (Premack, 1971, p. 148). Though in the previous paragraph of his article Premack had restated the importance of the probability-differential condition for determining reinforcement, in this one statement he abandoned the probability-differential hypothesis in favor of a condition of momentary deviation from baseline. Such a viewpoint is potentially congruent with the present molar equilibrium theory. It simply calls attention to its momentary aspects.

Mixed evidence for the role of momentary deviations from baseline probabilities was provided by Mazur (1977) who showed that the conditional probability of drinking but not running increased as a function of restricting access. However, both responses increased under appropriate deficit schedules. On the negative side, Bernstein and Dearborn (1978) reported a decrease in the conditional probability of the contingent response under the schedule when compared with its overall probability in the baseline (see Table VI). Yet, they obtained clear increases in the instrumental response. Based on these data, it seems that changes in the local probability of responding may accompany the imposition of a schedule, but overt changes in probability are not necessary to produce a change in instrumental responding. However, the further investigation of momentary disequilibria seems indicated (Allison & Timberlake, 1975; Ellis & Timberlake, 1977; Solomon & Corbit, 1974).

B. RESPONSE STRUCTURE

Dunham (1977) suggested that the essence of momentary probability could be captured by reference to changes in the molecular organization of behavior, namely, decreased burst length, and increased interburst interval of the contingent response. There is evidence that changes in burst length and interburst interval of contingent responding occur under schedules that produce an increase in the instrumental response (Premack, Schaeffer, & Hundt, 1964). However, there is no clear evidence that these changes are causal, rather than correlational, and there is good indirect evidence that such changes are not sufficient to produce an increase in instrumental responding. The data in previous sections that showed constant levels of responding under scheduled ratios made up of different sized terms suggest strongly that these particular changes in organization are not the key determinants of instrumental performance. Even those

TABLE VI

BASELINE PROBABILITIES AND MOMENTARY (CONDITIONAL) PROBABILITIES OF THE CONTINGENT RESPONSE[a]

Subject	Response	Baseline probability	Momentary (conditional) probability under schedule
MM	Read	.39	.20
ER	Read	.46	.23
GB	Talk	.40	.27
GW	Talk	.39	.29
CP	Yoga	.10	.05
CP	Reading	.07	.04
SL	Playing music	.13	.07
SL	Exercise	.03	.01
TN	Writing	.08	.02

[a] Assembled from Bernstein and Dearborn (1978).

experiments that showed changes in responding as a function of manipulating the sizes of the schedule terms, showed bidirectional changes in responding, rather than the unidirectional changes demanded by the hypothesis.

The response structure approach also suffers from problems of insufficient development (Podsakoff, 1980). If the molecular structure of the contingent response shows resistance to change, it would seem that the structure of instrumental responding should show similar properties (e.g., Schaeffer, 1966). A greater problem is to specify ahead of time the schedule conditions that will produce appropriate changes in burst-length and interburst interval. The actual values of these measures under a schedule are often under the control of the subject, serving more as dependent variables than as independent determinants. A more reasonable way to test the molecular equilibrium notions underlying Dunham's approach would be to impose schedules that manipulate the molecular characteristics of responding directly. For example, one could measure the baseline point of the instrumental and contingent responses in a behavior space defined by their interburst intervals. Then a schedule relating these measures could be imposed, and predictions made.

In summary, attempts to relate the molecular structure of behavior to molar schedule responding are important, and often intuitively appealing. However, to date they have suffered from an inexactness in the prior specification of the conditions for changes in responding, and lack of proof that they determine changes in behavior independently of their relation to molar equilibrium effects.

V. Summary

The study of learned performance has been dominated by a concern with the effects of associative variables. There has been too little concern with the functional role of learned behavior in promoting survival in the fact of environmental challenges. The functionalist tradition attempted to provide a regulatory basis for learned behavior by identifying substances, states, and set points important for survival. However, the classic approaches of Thorndike (1911) and drive theory proved inadequate to this task. The empirical approach identified reinforcers, but provided an incomplete description of the conditions surrounding their operation. Further, the empirical approach did little to advance our understanding of the regulatory processes determining behavior.

Premack's alternative to the empirical approach, the probability-differential hypothesis, was based on an important but incomplete analysis of the regulatory bases of reinforcement. The present molar equilibrium approach stems from Premack's work and employs paired baseline procedures to establish the set points of behavior. Within the equilibrium approach, reinforcers are not unique events but are created by restrictions imposed on behavior that prevent the subject from performing at its baseline set points.

Considerable evidence supports the predictions of the equilibrium approach concerning the conditions underlying learned performance. Neither the empirical law of effect nor the probability-differential hypothesis deals adequately with these data. Further, the equilibrium hypothesis predicts much of the data relating independent variables to changes in amount of instrumental responding, but it is inadequate in conditions of large initial disequilibrium and differential costs of deviations from the baselines of the instrumental and contingent responses.

A more complex equilibrium theory must consider: (*a*) the relative importance of the direction and amount of deviation of responding from baselines, (*b*) the effects of associative variables, (*c*) complex equilibria involving other responses inside and outside the session, and (*d*) the molecular and interactive determinants of response change.

From an equilibrium viewpoint, a schedule is a useful device for probing the operating characteristics of the organism. Behavior under schedules is the result of complex adjustments to conditions of disequilibrium imposed by the schedule. The potential generality of these adjustments holds out hope that a complete theory of behavior under schedules will account for much of behavior outside the schedule situation, not because the organism's behavior is based on learning, but because learning occurs in the context of the organism's behavior.

REFERENCES

Allison, J. Contrast, induction, facilitation, suppression and conservation. *Journal of the Experimental Analysis of Behavior,* 1976, **25**, 185–199.

Allison, J., Miller, M., & Wozny, M. Conservation in behavior, *Journal of Experimental Psychology: General,* 1979, **108**, 4–34.

Allison, J., & Timberlake, W. Instrumental and contingent saccharin licking in rats: Response deprivation and reinforcement. *Learning and Motivation,* 1974, **5**, 231–247.

Allison, J., & Timberlake, W. Response deprivation and instrumental performance in the controlled amount paradigm. *Learning and Motivation,* 1975, **6**, 122–142.

Angell, J. R. The province of functional psychology. *Psychological Review,* 1907, **14**, 61–91.

Bauermeister, J. J. Asymptotic reinforced responding as a function of the operant level of the instrumental response. *Learning and Motivation,* 1975, **6**, 143–155.

Bauermeister, J. J., & Schaeffer, R. W. Reinforcement relation: Reversibility within daily experimental sessions. *Bulletin of the Psychonomic Society,* 1974, **3**, 206–208.

Bentham, J. *An introduction to the principles of morals and legislation.* London: Payne, 1789.

Berlyne, D. E. *Conflict, arousal, and curiosity.* New York: McGraw-Hill, 1960.

Bernard, C. *Leçons sur les phenomènes de la vie en communs aux animaux et aux vegetaux.* Paris: J. Vrin, 1966. (Originally published, 1878).

Bernstein, D. J. *A comparison of cognitive and behavioral assessments of value in human behavior.* Paper presented at the meeting of the Midwestern Psychological Association, Chicago, 1979.

Bernstein, D. J., & Dearborn, M. *Utility of time-based theories of reinforcement in human behavior.* Unpublished manuscript, University of Nebraska, 1978.

Bernstein, D., & Ebbesen, E. B. Reinforcement and substitution in humans: A multiple response analysis. *Journal of the Experimental Analysis Behavior,* 1978, **30**, 243–254.

Bernstein, D. J., & Hinkle, P. H. *A re-examination of response deprivation in humans: A necessary control.* Unpublished manuscript, University of Nebraska, 1977.

Birch, D., Bernstein, E., & Clark, R. A. Response strength as a function of hours of food deprivation under a controlled maintenance schedule. *Journal of Comparative and Physiological Psychology,* 1958, **51**, 350–354.

Burkhard, B., Rachlin, H., & Schrader, S. Reinforcement and punishment in a closed system. *Learning and Motivation,* 1978, **9**, 392–410.

Campbell, B. A., Smith, N. F., Misanin, J. R., & Jaynes, J. Species differences in activity during hunger and thirst. *Journal of Comparative and Physiological Psychology,* 1966, **61**, 123–127.

Cannon, W. B. The physiological basis of thirst. *Proceedings of the Royal Society of London,* 1918, **90B**, 283–301.

Cannon, W. B. Hunger and thirst. In G. Murchison (Ed.), *Handbook of general experimental psychology.* Worcester: Clark University Press, 1934.

Carr, H. A. *Psychology, a study of mental activity.* New York: Longmans Green, 1925.

Collier, G. H., Hirsch, E., & Hamlin, P. H. The ecological determinants of reinforcement in the rat. *Physiology and Behavior,* 1972, **9**, 705–716.

Craig, W. Appetites and aversions as constituents of instincts. *Biological Bulletin,* 1918, **34**, 91–107.

Danaher, B. G. Theoretical foundations and clinical applications of the Premack principle: A review and critique. *Behavior Therapy,* 1974, **5**, 307–324.

Dashiell, J. F. *Fundamentals of general psychology.* Boston: Houghton Mifflin, 1928.

Digby, K. *Two treatises. In the one of which, the nature of bodies; in the other, the nature of mans soule; is looked into.* Paris, 1644.

Donahoe, J. W. Some implications of a relational principle of reinforcement. *Journal of the Experimental Analysis of Behavior,* 1977, **27**, 341–350.

Dufort, R. H., & Wright, J. H. Food intake as a function of duration of food deprivation. *Journal of Psychology,* 1962, **53,** 465-468.

Dunham, P. The nature of reinforcing stimuli. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior.* New York: Prentice-Hall, 1977.

Ebbesen, E. B., Allen, R., & Kjos, G. *Interaction of structure and function in response repertoires.* Paper presented at the meeting of the Midwestern Association for Behavior Analysis, Chicago, 1977.

Eisenberger, R., Karpman, M., & Trattner, J. What is the necessary and sufficient condition for reinforcement in the contingency situation? *Journal of Experimental Psychology,* 1967, **74,** 342-350.

Elliott, M. H. The effect of change of reward on the maze performance of rats. *University of California Publications in Psychology,* 1928, **4,** 19-30.

Ellis, R. W., & Timberlake, W. *A regulatory analysis of reward magnitude effects on alley running in rats.* Paper presented at the meeting of the Midwestern Psychological Association, Chicago, 1977.

Estes, W. K. Toward a statistical theory of learning. *Psychological Review,* 1950, **57,** 94-107.

Falk, J. L. The nature and determinants of adjunctive behavior. *Physiology and Behavior,* 1971, **6,** 577-588.

Fiske, D. W., & Maddi, S. R. A conceptual framework. In D. W. Fiske & S. R. Maddi (Eds.), *Functions of varied experience.* Homewood, Ill.: Dorsey, 1960.

Fitzsimmons, J. T. Thirst. *Physiological Review,* 1972, **52,** 468-561.

Guthrie, E. R. *The psychology of learning.* New York: Harper, 1935.

Harlow, H. F., Harlow, M. K., & Meyer, D. R. Learning motivated by a manipulation drive. *Journal of Experimental Psychology,* 1950, **40,** 228-234.

Harrison, R. G., & Schaeffer, R. W. Another test of the Premack principle. *Bulletin of the Psychonomic Society,* 1975, **6,** 565-568.

Haverland, E. M. *The application of an analytical solution for proportional profiles rotation to a box problem and to the drive structure of rats.* Unpublished doctoral dissertation, University of Illinois, 1954.

Hearst, E., & Jenkins, H. M. *Sign tracking: The stimulus-reinforcer relation and directed action.* Austin, Texas: The Psychonomic Society, 1974.

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

Heth, C. D., & Warren, A. G. Response deprivation and response satiation as determinants of instrumental performance: Some data and theory. *Animal Learning & Behavior,* 1978, **6,** 294-300.

Hilgard, E. R. *Theories of learning* (2nd ed.). New York: Appleton, 1948.

Hilgard, E. R., & Bower, G. H. *Theories of learning* (4th ed.). New York: Prentice-Hall, 1975.

Hogan, J. A., & Roper, T. J. A comparison of the properties of different reinforcers. In J. S. Rosenblatt, R. A. Hinde, C. Beer, & M.-C. Busnel (Eds.), *Advances in the study of behavior* (Vol. 8). New York: Academic Press, 1978.

Holstein, S. B., & Hundt, A. G. Reinforcement of intracranial self stimulation by licking. *Psychonomic Science,* 1965, **3,** 17-18.

Hull, C. L. Special review: Thorndike's fundamentals of learning. *Psychological Bulletin,* 1935, **32,** 807-823.

Hull, C. L. *Principles of behavior.* New York: Appleton, 1943.

Iverson, I. H. Interactions between reinforced responses and collateral responses. *The Psychological Record,* 1976, **26,** 399-413.

Jacobson, E., & Premack, D. Choice and habituation as measures of response similarity. *Journal of Experimental Psychology,* 1970, **85,** 30-35.

Jennings, H. S. *The behavior of lower organisms.* New York: Columbia University Press, 1906.

Kanarek, R. B. Energetics of meal patterns in rats. *Physiology and Behavior,* 1976, **17,** 395–399.

Kavanau, J. L. Behavior of captive white-footed mice. In E. P. Willems & H. L. Raush (Eds.), *Naturalistic viewpoints in psychological Research.* New York: Holt, 1969.

Kelsey, J. E., & Allison, J. Fixed-ratio level pressing by VMH rats: Work vs accessibility of sucrose reward. *Physiology and Behavior,* 1976, **17,** 749–754.

Kjos, G. L. *Constraints on the reinforcement relation: Tests of instrumental response type in a multiple response repertoire environment.* Unpublished doctoral dissertaion, University of California, San Diego, 1977.

Klajner, F. *The relations among instrumental performance, reinforcement, and contingent-response deprivation in the instrumental conditioning paradigm.* Unpublisbed doctoral dissertation, University of Toronto, 1975.

Konarski, E. A., Jr., *The necessary and sufficient conditions for increasing instrumental responding in the classroom: Response deprivation vs. probability differential.* Unpublished doctoral dissertation, University of Notre Dame, 1979.

Konarski, E. A., Jr., Johnson, M. R., Crowell, C., & Whitman, T. L. *Response deprivation, reinforcement, and instrumental academic performance in an EMR classroom.* Paper presented at the thirteenth annual Gatlinburg Conference on Research in Mental Retardation and Developmental Disabilities, Gatlinburg, Tennessee, 1980.

Lepper, M. R., & Greene, D. *The hidden costs of reward: New perspectives on the psychology of human motivation.* Hillsdale, N. J.: Erlbaum, 1978.

Lewin, K. Field theory and learning. In *The psychology of learning.* 41st Yearbook of the National Society for the Study of Education, Part II, 1942, pp. 215–242.

Lockard, R. B. A method of analysis and classification of repetitive response systems. *Psychological Review,* 1964, **71,** 141–147.

Loeb, J. *Forced movements, tropisms, and animal conduct.* Philadelphia: Lippincott, 1918.

Luce, R. D. *Individual choice behavior: A theoretical analysis.* New York: Wiley, 1959.

McFarland, D. J. Decision making in animals. *Nature (London),* 1977, **269,** 15–21.

McIntosh, J. W. *Response deprivation theory of instrumental responding tested with a freely available response: A motivational interpretation.* Unpublisbed doctoral dissertation, Indiana University, 1974.

Marmaroff, S. *Reinforcement: A test of Premack's differential probability rules.* Unpublished master's thesis, Dalhousie University, 1971.

Marwine, A., & Collier, G. The rat at the waterhole. *Journal of Comparative and Physiological Psychology,* 1979, **93,** 391–402.

Mazur, J. E. The matching law and quantifications related to Premack's principle. *Jouranl of Experimental Psychology: Animal Behavior Processes,* 1975, **104,** 374–386.

Mazur, J. E. Quantitative studies of reinforcement relativity. *Journal of the Experimental Analysis of Behavior,* 1977, **27,** 137–150.

Meehl, P. E. On the circularity of the law of effect. *Psychological Bulletin,* 1950, **45,** 52–75.

Miller, N. E. Shortcomings of food consumption as a measure of hunger: Results from other behavioral techniques. *Annals of the New York Academy of Sciences,* 1955, **63,** 141–143.

Pavlov, I. P. *Lectures on conditioned reflexes.* (Translated by W. H. Gantt.) New York: International Publishers, 1928.

Peden, B., & Timberlake, W. *Fixed-ratio and variable-interval keypecking: Effects of reward magnitude.* Paper presented at the meeting of the American Psychological Association, Washington, D.C., 1976.

Podsakoff, P. M. *Performance models, microeconomics, and schedule behavior in humans.* Unpublished doctoral dissertation, Indiana University, 1980.

Powell, R. W., & Curley, M. Instinctive drift in nondomesticated rodents. *Bulletin of the Psychonomic Society,* 1978, **18,** 175–178.

Premack, D. Toward empirical behavior laws. I. Positive reinforcement. *Psychological Review,* 1959, **66,** 219-234.

Premack, D. Rate differential reinforcement in monkey manipulation. *Journal of the Experimental Analysis of Behavior,* 1963, **6,** 81-89.

Premack, D. Reinforcement theory. In D. Levine (Ed.), *Nebraska symposium on motivation, 1965.* Lincoln: University of Nebraska Press, 1965.

Premack, D. Catching up with common sense or two sides of a generalization: Reinforcement and punishment. In R. Glaser (Ed.), *The nature of reinforcement.* New York: Academic Press, 1971.

Premack, D., & Collier, G. Analysis of nonreinforcement variables affecting response probability. *Psychological Monographs,* 1962, **80,** (5, Whole No. 524).

Premack, D., Schaeffer, R. W., & Hundt, A. Reinforcement of drinking by running: Effect of fixed ratio and reinforcement time. *Journal of the Experimental Analysis of Behavior,* 1964, **7,** 91-96.

Rachlin, H., & Burkhard, B. The temporal triangle: Response substitution in instrumental conditioning. *Psychological Review,* 1978, **85,** 22-45.

Richter, C. P. A behavioristic study of the activity of the rat. *Comparative Psychology Monographs,* 1922, **1**(2).

Richter, C. P. Animal behavior and internal drives. *Quarterly Review of Biology,* 1927., **2,** 307-343.

Richter, C. P. Total self-regulatory functions in animals and human beings. *Harvey Lectures,* 1942-1943, **38,** 63-103.

Roper, T. J. Nesting material as a reinforcer for female mice. *Animal Behaviour,* 1973, **21,** 733-740.

Roper, T. J. Nest material and food as reinforcers for fixed-ratio responding in mice. *Learning and Motivation,* 1975, **6,** 327-343.

Schaeffer, R. W. The reinforcement relation as a function of instrumental response base rate. *Journal of Experimental Psychology,* 1965, **69,** 419-425.

Schaeffer, R. W. Overshooting of the FR requirement. *Psychological Record,* 1966, **16,** 17-23.

Sheffield, F. D., Wulff, J. J., & Backer, R. Reward value of copulation without sex drive reduction. *Journal of Comparative and Physiological Psychology,* 1951, **44,** 3-8.

Shettleworth, S. J. Constraints on learning. In D. S. Lehrman, R. A. Hinde, & E. Shaw (Eds.), *Advances in the study of behavior* (Vol. 4). New York: Academic Press, 1972.

Skinner, B. F. *The behavior of organisms: An experimental analysis.* New York: Appleton, 1938.

Solomon, R. L., & Corbit, J. D. An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychological Review,* 1974, **81,** 119-145.

Spencer, H. *Principles of psychology.* New York: Appleton, 1883.

Staddon, J. E. R. Operant behavior as adaptation to constraint. *Journal of Experimental Psychology: General,* 1979, **108,** 48-67.

Terhune, J. G. The relationship between momentary response probabilities and momentary reinforcement effects. *Animal Learning and Behavior,* 1978, **6,** 187-192.

Thorndike, E. L. *Animal intelligence.* New York: Macmillan, 1911.

Thorndike, E. L. *The Psychology of Learning.* (Educational Psychology, II.) New York: Teachers College, Columbia University, 1913.

Timberlake, W. *Fixed-ratio schedules of wheel running and saccharin licking in rats: Response deprivation and instrumental performance.* Unpublished manuscript, Indiana University, 1974.

Timberlake, W. The application of the matching law to simple ratio schedules. *Journal of the Experimental Analysis of Behavior,* 1977, **25,** 215-217.

Timberlake, W. Licking one saccarin solution for access to another in rats: Contingent and noncontingent effects in instrumental performance. *Animal Learning and Behavior,* 1979, **7,** 277-288.

Timberlake, W., & Allison, J. Response deprivation: An empirical approach to instrumental performance. *Psychological Review,* 1974, **81,** 146-164.

Timberlake, W., & Peden, B. *Keypecking for food in pigeons: Regulation or reinforcement?* Paper presented at meeting of the Midwestern Association for Behavior Analysis, Chicago, 1977.

Timberlake, W., & Wozny, M. Reversibility of reinforcement between eating and running by schedule changes: A comparison of hypotheses and models. *Animal Learning and Behavior*, 1979, **7**, 461–469.

Walker, E. L. Psychological complexity as a basis for a theory of motivation and choice. In D. Levine (Ed.), *Nebraska symposium on motivation, 1964*. Lincoln: University of Nebraska Press, 1964.

Warden, C. J., Jenkins, T. N., & Warner, L. H. *Comparative psychology*. New York: Ronald, 1936.

Wasik, B. H. The effects of fixed ratio and contingent time on human lever-pressing behavior. *Psychological Record*, 1969, **19**, 95–104.

Woodworth, R. S. *Dynamic Psychology*. New York: Columbia University Press, 1918.

Wozny, M. *Models and microeconomics of performance under fixed-ratio schedules*. Unpublished doctoral dissertation, Indiana University, 1979.

FISH AS A NATURAL CATEGORY FOR PEOPLE AND PIGEONS[1]

R. J. Herrnstein

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS

Peter A. de Villiers

SMITH COLLEGE
NORTHAMPTON, MASSACHUSETTS

---

[1] We gratefully acknowledge James Q. Wilson's generosity and trust in allowing us to reproduce about 350 of his 35-mm slides of underwater scenes. Wilson used a Nikon underwater SLR camera and took most of the pictures using artificial light, although a few, nearer the water surface, used natural light. Special thanks are also owed to Jill de Villiers for commenting on earlier drafts of the paper and for participating in the initial planning of the study. Erica Paquette deserves thanks, too, for volunteering to run subjects when help was short. Preparation of the article and some of the laboratory facilities were supported by Grant MH-15494 from NIMH to Harvard University. Finally, we must thank the William F. Milton Fund of Harvard University, which paid for the reproduction of the slide collection.

THE PSYCHOLOGY OF LEARNING
AND MOTIVATION, VOL. 14

ISBN 0-12-543314-X

Epigraph

Letter III from The Unnatural History of Selborne

According to *The Observer's Book of Trees* just purchased, what I had been leaning on was a stick hewn from the common juniper. This is odd, since the common juniper is a bush, whereas ours is ninety feet high and full of spoonbills. Or possibly bats. It may therefore be an uncommon juniper, which does not figure in the index. . . . My small daughter afforded me much relaxed laughter by insisting that the tree was a larch! I reminded her that the larch was a fish, but not sternly; she is but six, and easily upset by scorn.

I now know . . . that the animal I observed as I fell was neither fox nor dog, but a weasel. I identified it instantly from my new *Observer's Book of Wild Animals*. Ours is a large specimen, about the size of a child's tricycle, and I have advised the family to keep well away from it.

This evening, my son and I embarked upon a pleasant excursion to collect examples of the wild flowers with which this part of the forest is so abundantly blessed. We collected a daisy, and fifty-nine things that weren't. (Coren, 1979, p. 962)

I. Introduction

In several recent experiments, pigeons, blue jays, and primates have learned to classify photographic stimuli exemplifying open-ended categories and have generalized to new instances of the category without difficulty (reviewed in Herrnstein, 1979). Pigeons, for example, have learned to classify photographs of trees, bodies of water, other pigeons, oak leaves, people in general or an individual person, and alphanumeric characters; blue jays have classified photographs of cryptic moths; apes and monkeys have classified photographs of other apes or monkeys. Cerella (1979) demonstrated that pigeons trained to distinguish between a single leaf of the white oak and various examples of leaves from other deciduous trees spontaneously generalized to other leaves from the white oak. Although oak leaves obviously resemble each other, it is not obvious how to specify the resemblance in physical terms that could distinguish them objectively from, say, maple leaves, tulip-tree leaves, elm leaves, etc. For the reader whose memory of leaf-forms does not support this claim, we provide Fig. 1, illustrating six typical stimuli used in Cerella's study. Silhouettes of leaves of approximately constant vertical extent and in a fixed orientation were either from the white oak (top row) or not. Training with just the center oak leaf versus a variety of non-oaks led to spontaneous correct generalization to oak leaves as different as those on

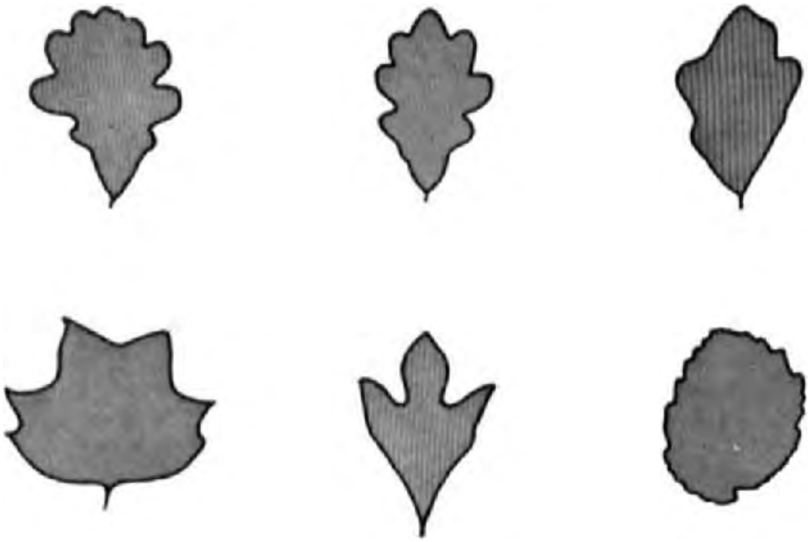


Fig. 1. Showing six of the stimuli used by Cerella (1979) in a study of white oak leaf discrimination by pigeons. The three upper leaves are from white oak; the bottom three are not. Copyright (1979) by the American Psychological Association. Reprinted by permission.

either side. The class rule used by the pigeons did not seem to rely on the obvious metrical features of leaves, such as number of lobes, depth of lobes, or the like. It is, in fact, much easier to intuit the pigeons' rules, which no doubt involved form, than to state them.

Cerella also showed that it was much harder for the pigeons to learn to distinguish an individual oak leaf from other oak leaves than from nonoak leaves, using a successive discrimination procedure in which only a single leaf is present at any moment. What is retained in memory about an oak leaf apparently encompasses other oak leaf-like forms, making a discrimination among leaves from within a species of tree harder in this instance than one between species. From an evolutionary standpoint, the benefits of recognition of classes at some taxonomic level—e.g., species or genera—are probably more common and more pressing than recognition of specific instances. Pigeons, for example, need to recognize species of seeds or predators more than they need to recognize given kernels of corn or given cats. Perceptual systems may therefore tend to evolve so as to be "transparent," to use Cerella's apt word, to the kinds of variation within the functionally relevant taxa. Even for recognizing the unique individual, transparent variation helps, for the individual object probably appears in differing perspectives and differing conditions of illumination on different occasions. Variation is thus the rule for the classes of stimuli that organisms must deal with, and so we should be prepared to find complexity, subtlety, and power in the ability to categorize even by lower organisms, once we use naturalistic stimuli in experiments.

Trees or people as stimuli for pigeons and moths as stimuli for blue jays seem naturalistic in the sense that they are probably important functional classes in any pigeon's or blue jay's natural habitat. This has generally been the case in the research on categorizations of naturalistic stimuli. In the experiments to be described here, we therefore used naturalistic stimuli not drawn from the pigeon's natural environment. We procured several hundred 35-mm slides of underwater scenes taken by a recreational Scuba diver during vacation trips to the Caribbean and tropical Pacific. None of the pictures were taken for use in experiments. In all but one of the experiments, the subjects were required to discriminate pictures containing fish from those not containing fish. In some of the pictures the fish was apparently the object of interest, but many pictures contained fish only incidentally. In about 40% of the slides containing fish more than one fish was visible. Generally, the pictures are typical of tropical waters, including vegetation, coral, various animals in addition to fish, rock formations, and occasional debris of shipwrecks. The fish varied from fairly typical examplars such as tuna or barracuda to more exotic varieties such as stonefish and more doubtful instances like eels and skates. Some 30% of non-fish slides contained other creatures such as snails, jelly fish, lobsters, crabs, starfish, and turtles.

The ancestors of pigeons have not shared an environment with the ancestors of fish for about 75 million years, at which time fish may have looked rather different from the ones in our pictures. Consequently, it is a priori unlikely that anything approximating a fish category could be genetically programmed in our subjects and the data make it more unlikely still, as we will explain later. It is also unlikely that our pigeons in particular ever saw a fish, although some pigeons, living near water, may occasionally see a fish washed ashore. However, to our knowledge, pigeons are genuinely herbivorous and would have no direct interest in fish as such. In short, neither the genetic nor the experiential sources would seem to predispose pigeons to recognize fish as a class, at least in any direct way. If the class forms anyway, we must consider more indirect predispositions.

The plan of the article is to describe and compare the results of three experiments, one on human subjects and two on pigeons, and to draw whatever conclusions we may about categorization in light of the results.

II. Experiment 1: Human Subjects

The purpose of this experiment was to establish rankings of the photographs for human subjects, using reaction time to quantify the discriminability or acceptability of the fish or non-fish in a collection of 35-mm slides. Eighty slides containing single instances of fish and 80 slides covering the range of views without fish were selected. The same slides were subsequently used in the pigeon experiments.

A. METHODS

1. Subjects

Twelve students (undergraduate and graduate) volunteered to serve in a single session lasting about 30 min.

2. Apparatus

The subject was seated in a chair approximately .75 m from a rear-projected screen. Slides were projected at eye level so that they subtended an angle of 15° at the eye. In front of the subject were two telegraph keys, one for each hand.

3. Procedure

The subjects were instructed to depress one telegraph key if they saw any part of a fish in a picture and to depress the other if they did not. Right- and left-hand keys were alternated across subjects without regard to handedness. One hundred and sixty 35-mm slides were presented. Between each slide there was a 5-sec intertrial interval, followed by the presentation of a fixation point for 2 sec before the slide came on. The fixation point consisted of a bright spot of light on a dark grey field and the slide was presented centered on the dot. Half of the slides contained fish and half did not, as far as the experimenters could determine (with help from the photographer). The pictures were randomized anew for each subject. Latencies were recorded in hundreths of a second, from the onset of each picture to the operation of one of the keys.

B. RESULTS AND DISCUSSION

Figure 2 contains frequency distributions of mean reaction times (RT) over the 12 subjects for the 80 slides containing fish (positives) and the 80

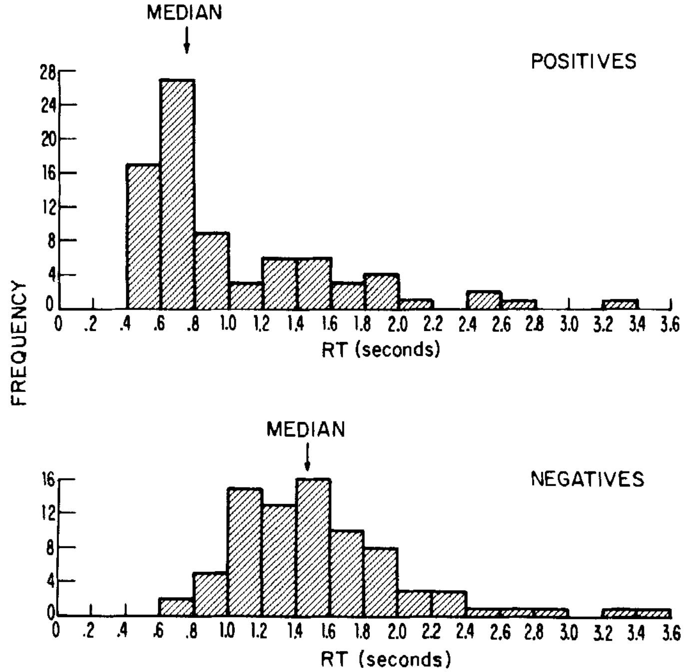


Fig. 2. Human reaction times (RT) tallied in class intervals of .2 sec. Twelve subjects saw, in different random orders, 80 photographic slides containing fish ("positives") and 80 not containing fish ("negatives") and pressed one button or another to indicate whether or not they saw a fish in each photograph. RT was measured from the presentation of a slide to the depression of a button.

containing no fish (negatives), in class intervals of .2 sec. The median RT for positives (.76 sec) was about half that for negatives (1.47 sec). The standard deviations (in seconds) of the individual RT for positive (σ_p) and negative (σ_n) stimuli were reasonably well described by the following best-fitting regression lines:

$$\sigma_p = .959(\mathrm{RT})_p - .355 \qquad (.83) \qquad (1a)$$

$$\sigma_n = 1.07(\mathrm{RT})_n - .488 \qquad (.79) \qquad (1b)$$

The decimals in parentheses give the proportions of variance accounted for by the fitted equations.

The accuracy of the discrimination was generally high, especially for the negative instances. For negatives, 54 of the 80 stimuli had zero errors over all subjects and the total number of misclassifications was 39 out of a possible 960 (12 subjects times 80 stimuli). For positives, 38 out of the 80 stimuli had zero errors and the total number of misclassifications was 128 out of 960 opportunities. No negative or positive stimulus was incorrectly classified by more than 6 out of the 12 subjects.

Misclassifications were too scarce to provide a reference scale of difficulty in classification, which was the main point in running the human subjects. However, RT provided broad ranges of values for positives and negatives, and Fig. 3 indicates that they were closely associated with errors. RT was averaged for stimuli sorted according to the number of misclassifications by the 12 subjects. Points are plotted only if there were at least four stimuli contributing to the average, which yielded six data points for positives and three for negatives. RT was consistently about twice as long for negatives as positives, but each function rises unmistakably. The more ambiguous the stimulus, as measured by misclassifications, the larger the average RT across subjects. We concluded that we may therefore use RT as a rough index of difficulty in classification as a basis of comparison with the results from pigeons.

In fact, the indices we use in those comparisons are rankings from 1 to 80 (for positives and negatives separately) of the mean rank earned by each stimulus when each subject's RT was ranked separately. If all subjects ranked the stimuli identically then the mean ranks would themselves have run from 1 to 80. To the extent that the subjects disagreed, the mean ranks would tend to regress toward intermediate values, centering on 40.5. It is therefore of some interest to calculate the actual degree of regression of the mean ranks, as a measure of intersubject agreement.

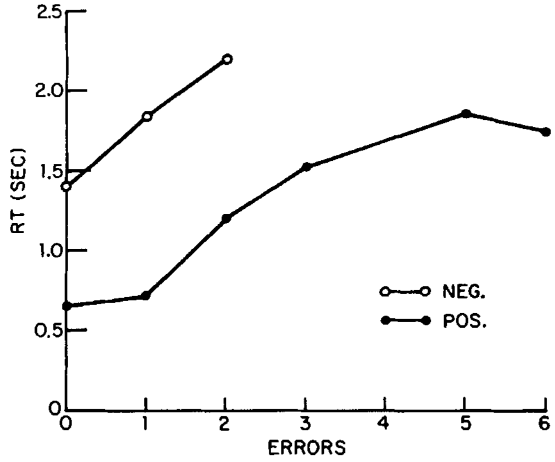


Fig. 3. Human reaction time as a function of the number of times a slide was misclassified by the 12 subjects. Each subject saw the 160 slides once. Upper line is for slides not containing fish; lower line, for those containing fish.

Regressing the mean ranks ($\bar{r}$) on the rank of the mean rank ($\bar{R}$), for positives and negatives, gives:

$$\bar{r}_p = .809\ R_p + 7.79 \qquad (.98) \qquad (2a)$$

$$\bar{r}_n = .566\ R_n + 17.72 \qquad (.97) \qquad (2b)$$

The decimals in parentheses give proportions of variance accounted for by the regression equations. If the subjects had agreed perfectly in their ranked RTs, the slopes would have been 1.0 and the intercepts, 0. There was greater agreement among subjects for positives [Eq. (2a)] than for negatives [Eq. (2b)]. This probably reflects a greater ambiguity in the set of positive stimuli, because intersubject agreement about difficulty implicitly assumes a range of difficulty. The intersubject agreement is lower among negatives probably because the negative set was more homogeneous (i.e., homogeneously easy), as the data on misclassifications demonstrated. It took the subjects longer to classify a picture as negative, but in general they were more often correct when they did so.

We examined the 160 pictures ourselves, with the results from the 12 subjects in hand. We were looking for those characteristics of pictures that seemed to contribute to RT. Three main factors seemed to be operating, as follows:

1. Contrast or figure-ground. The pictures varied in the visibility of elements that may or may not have been fish. A picture with a single figure on a homogeneous ground typically produced a short RT, whether positive or negative. "Busy" pictures typically produced long RTs.
2. Prototypicality. Some species of fish look more like people's idea of fish than others do (see Rosch, 1978, on prototypicality more generally). Odd-looking fish, such as the eels, skates, stonefish, and gars in the set, tended to have longer RT. In principle, there could be considerations of prototypicality for non-fish, but our data did not illustrate them in any obvious way.
3. Canonical view. The shortest RT for positives were to side views of fish (assuming prototypicality and clear figure-ground contrast), especially if the animal's eye was clearly visible.

We have not attempted to quantify these three factors, mainly because a sample of only 160 photographs can provide relatively scanty evidence for them at best. They are offered here as a rough sketch of the controlling variables, which we will elaborate further in Section V, in light of the results of all three experiments.

C. SUMMARY

Twelve human subjects responded to 160 underwater photographs to indicate whether they saw a fish or not. Judging from the small number of errors in classification, this was a relatively easy task, but a range of reaction times suggested that some pictures were easier to classify than others. With reaction time as an index of difficulty or ambiguity, the subjects ranked the positive instances more similarly than the negative instances. Both positive and negative instances were ranked similarly enough to allow a meaningful composite ranking over all subjects, as a basis for comparison with rankings by pigeons.

III. Experiment 2: Discrimination and Generalization of Fish by Pigeons

Pigeons were reinforced for pecking in the presence of underwater photographs containing fish and not reinforced in the presence of underwater photographs containing no fish. Every session used a new set of photographs. In addition, there were generalization tests using stimuli from Experiment 1. The purpose of the experiment was to see if pigeons can induce a category approximating fish from a large sample of instances seen once or infrequently, and to compare rankings by pigeons with rankings by people.

A. METHOD

1. Subjects

Subjects were four adult male white Carneaux pigeons at 80% of *ad lib* weight. The pigeons had worked in a previous experiment in which they learned to peck in the presence of photographs containing trees. Their first encounter with underwater photographs was the first session in the present experiment.

2. Apparatus

A standard operant chamber for pigeons was equipped with a translucent screen, measuring .044 by .064 m, in one wall. The screen was hinged so that a peck of .20 N operated a switch that delivered an electrical pulse. A Kodak Carousel projector was mounted behind the screen, adapted so that it could be advanced remotely by the computer (PDP8) that controlled experimental sessions and recorded the data. Food reinforcements were presented by a hopper mounted in the wall beneath the screen.

3. *Procedure*

In a session, 80 slides were presented for an average of 30 sec each, varying between 10 and 50 sec. Positive slides, containing fish, were associated with a variable-interval schedule of food reinforcement averaging 45 sec for pecks directly at the screen. The food hopper operated for 5 sec per reinforcement. Negative slides, containing no fish, were associated with neither reinforcement nor any penalty for pecking. During an intertrial interval of 5 sec the screen was dark but the houselight in the chamber remained on.

The 80 slides (40 positives and 40 negatives) in each session were drawn at random, with replacement, from a set of 120 positives and 120 negatives that included half of the 160 pictures used in Experiment 1. The other half was reserved for generalization tests as described below. No effort was made to restrict or otherwise select the sample of pictures shown in a session. The 80 stimuli were presented to all pigeons in a different random order, and new random orders were used in every successive session. The restrictions on randomness were that the 80 stimuli be half positive and half negative and that there be no more than four positives or negatives in a row.

The reinforcement schedule yielded zero reinforcements for about half of the positive stimuli. A small proportion of positives permitted more than a single reinforcement. The analysis was based on ranked rates of pecking, for which purpose we used only responding up to the first reinforcement in a slide's presence. The total presentation time could therefore be used for positives earning no reinforcement. The same VI 45-sec schedule was used for ascertaining the time base on which to compute the rate of pecking to negative slides, but of course no reinforcements were given while they were present.

The foregoing procedure, the "training" phase, was in force for 12 sessions, then there were five sessions of generalization "probes." In probe sessions, 64 stimuli (32 positive and 32 negative) were chosen as before from the set of 240, and the same contingencies were maintained. Because the sampling of stimuli was with replacement, some of these stimuli had been shown before. However, the remainder comprised 16 slides (8 positive, 8 negative) from the set shown to the human subjects and held in reserve. Consequently, the pigeons had seen none of these probe stimuli before. They appeared in a single session, randomized along with the other 64 slides, but the probes were always shown for exactly 30 sec and there were no reinforcements for pecking to either positive or negative probes.

Five probe sessions were run, each using a new set of 16 probe stimuli and a new set of 64 background stimuli. The probes were selected to span the range of difficulty for the human subjects. The 80 positives and 80 negatives from Experiment 1 were each subdivided into decades of mean rank of reaction times. Then, in each set of 16 probes, there was one stimulus from each decade. We chose mainly probes for which the between-human-subject variability in rank was small, except when this eliminated certain interesting cases, such as negative probes containing a turtle.

The final session immediately followed the five generalization probe sessions, so that there were 18 sessions altogether in Experiment 2. The final session used only the probe stimuli from the preceding five sessions. These 80 stimuli (40 positives and 40 negatives) had therefore been seen once before, but no reinforcements had been given in their presence. The procedure was as described above for the training phase, except that no reinforcements were permitted during the first 10 sec of any positive stimulus, thus providing at least 10 sec of usable data for each slide.

B. RESULTS AND DISCUSSION

Performance is conveniently summarized by the mean rank assigned to positive stimuli (r_p), based on ranked rates of pecking to the 80 stimuli in a session. Various relationships involving mean ranks should be noted first. The mean rank for a session is 40.5; the mean positive rank plus the mean negative rank must equal 81. The mean positive rank (or the mean negative rank) is linearly related to the Mann–Whitney U, as follows:

$$r_p = n + \frac{n+1}{2} - \frac{U}{n} \tag{3}$$

In this, and subsequent, equations, n is the number of positives or negatives for a complete session, namely, 40. Since probability levels for U are available, Eq. (3) enables us to assign probability levels[2] for values of

---

[2] This ignores the effects of ties on the sampling distribution of U, which would have been negligible here. As long as there is any responding at all in the presence of a stimulus, the chances of its being tied can be minimized by measuring rate to several (e.g., 6) significant figures in responses per .1 sec. However, once a discrimination forms, a number of stimuli are likely to tie at zero responding. One way to handle this (Herrnstein, 1979) is to have the computer rank across tied stimuli in order of their presentation in a session, which is to say, to rank them according to the random sequence in which they had been shown. Eliminating tied ranks this way simplifies data analysis, but it may not be appropriate for deeper levels of analysis than have been attempted yet.

r_p. When r_p equals 36.2, the associated value of U reaches a .05 level of statistical significance. When r_p is 30.0, the associated U is statistically significant at a probability level beyond .00003.

Initial acquisition for the four pigeons is shown in Fig. 4 in terms of r_p. Within the cross-hatched region, no statistically valid discrimination is being shown. The most rapid learning was by P3, whose first session yielded a significantly elevated value of r_p. About halfway through the first session, P3 began to sort with some accuracy. P2 and P4, in contrast to P3, were not discriminating significantly until the fifth session. P1 was intermediate, reaching statistical significance on the second session. In Herrnstein (1979), four pigeons learning to sort for the category of trees reached the .05 level in two to three sessions.

The horizontal lines in Fig. 4 are at the medians of the final five training sessions for each subject, which averaged 28.3 for the four subjects. This can be converted into a more meaningful number by taking advantage of another relationship involving r_p. Earlier papers (Herrnstein, Loveland, & Cable, 1976; Herrnstein, 1979) used an index of discrimination, ρ, which estimates the probability that a randomly chosen positive stimulus was ranked higher than a randomly chosen negative stimulus. In the absence of discrimination, ρ approximates .5. It has been shown (Bamber, 1975) that, for ranked data, ρ approximates the area under an ROC graph for "hits" versus "false alarms." The discrimination index is related to r_p by the following equation:

$$r_p = n + \frac{n+1}{2} - n\rho \qquad (4)$$

When r_p is 28.3, $\rho = .81$, which compares favorably with other data on pigeons. In Herrnstein *et al.* (1976) ρ averaged .85 for pigeons sorting the category of trees, .79 for bodies of water, and .79 for a particular human.

After the 12 training sessions came the five sessions of generalization probes, in which each set of 80 stimuli included 16 (8 fish and 8 non-fish) unreinforced, 30-sec presentations of stimuli being shown for the first time. The probe stimuli were drawn from the set used in Experiment 1 and each 16 spanned the range of difficulty for the human subjects, as measured by mean rank of reaction time. The results of these tests are summarized in Table I in terms of ρ and the concordance among the subjects.

The 16 probes in each session were ranked separately from the 64 other stimuli for Table I. Each set of 16 was significantly discriminated by at least two subjects; set 2 and 3 were significantly discriminated by all four subjects. Each subject discriminated significantly with either three or four

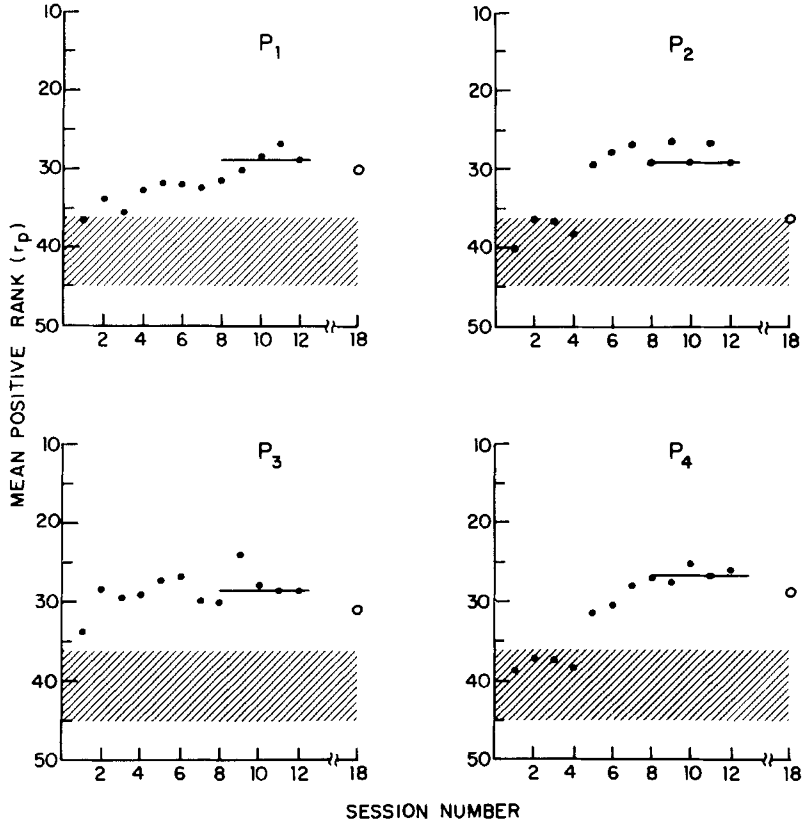


Fig. 4. Initial discrimination by the four pigeons in Experiment 2. In every session, subjects saw 80 different stimuli, 40 containing fish and 40 not containing any. Rates of responding in the presence of the stimuli were ranked from 1 (highest) to 80 (lowest), and the points show the mean rank for stimuli containing fish (r_p). Significant discrimination is being shown when a point is outside the cross-hatched region. The horizontal line is at the median r_p for sessions 8–12. Open point is for a generalization test, explained in text.

of the five sets of probes. Though there is clear evidence of a generalized discrimination, the probe stimuli varied in discriminability upon first viewing; the average value of ρ in Table I is .74, corresponding to r_p of 30.9, which is somewhat less of a discrimination than reached by the end of the training phase.

The concordances (Kendall's W) at the bottom of Table I further substantiate some sort of generalized category. This measure in effect averages the Spearman rank-order correlations between all pairs of subjects. Coefficients of concordance were calculated for the ranked rates of pecking to positive and negative probes separately, so as to exclude the con-

TABLE I

DISCRIMINATION (p) AND CONCORDANCE (W)
FOR FIVE SETS OF TEST STIMULI

	Discrimination: p				
Subjects	Set 1	Set 2	Set 3	Set 4	Set 5
P1	.422	.891***	.859**	.734	.750*
P2	.844**	.781*	.797*	.375	.750*
P3	.594	.828**	.813*	.797*	.609
P4	.750*	.938***	.875***	.750*	.719
	Concordance: W				
Positive stimuli	.554*	.224	.677**	.147	.489
Negative stimuli	.532*	.845***	.344	.483	.576*

*$.01 < p \leq .05$.
**$.005 < p \leq .01$.
***$p \leq .005$.

tribution of discrimination between positives and negatives. Also, it should be recalled that each subject saw the stimuli in a different random order, so that agreement is independent (or in spite) of the factor of order of appearance. Of the 10 tests of concordance in Table I, five achieved conventional statistical significance, twice for positive probes and three times for negative.

Other aspects of generalization are displayed in Fig. 5, which compares probe and baseline stimuli directly and also contrasts positive and negative stimuli. All the stimuli in each probe session were placed in a single ranking, then the ranks were subdivided for positive and negative, probe and baseline stimuli. These four subdivisions were then averaged across the five probe sessions for each subject and are plotted on Fig. 5. Baseline ranks are along the x-axis; probe-stimulus ranks, along the y-axis. Positive stimuli are open symbols; negative stimuli are closed symbols. Each subject has a different symbol, and the dashed line connects their averages. The solid line shows equal abscissa and ordinate values.

If baseline and probe ranks had been equal, the data would have lain along the solid line. In fact, both positive and negative probe ranks tended to fall closer to the mid-rank than the baseline stimulus ranks. This indicates a degree of generalization decrement, in that the probe stimuli were less sharply discriminated than the baseline stimuli even within the same sessions. However, all subjects showed less decrement for fish probes than for non-fish probes. What this means may have something to do with an earlier finding. Trees also generalize better than non-trees (Herrnstein, 1979) and they do so whether the trees are the positive or the negative category. If, as seems reasonable, fish or tree stimuli are being sampled from a smaller region in a multidimensional sensory space than non-fish or non-tree stimuli, we would expect less generalization decrement for trees or fish than for their absence.

Next, the 80 probe stimuli were combined in a single, final session, which is shown on Fig. 4 as session 18. Three of the four subjects displayed clearly significant discriminations, but all four showed somewhat less discrimination than they did at the end of the training phase, corroborating the generalization decrement in Fig. 5. The average value of r_p for the eighteenth session was 31.7, corresponding to a value of p equal to .72. This matches quite well the average of the probes in Table I, i.e., .74, both below the .81 reached at the end of training. It is possible that the probe stimuli were, on average, harder to classify than the typical random selection used in training, since the probes were explicitly chosen to span the range of difficulty for the human subjects. We will reconsider the relevance of the human rankings in Section V.

For determining the concordance in session 18, separate rankings for positive and negative stimuli were again used. For positives, Kendall's W was .513 and for negatives, it was .544. Both values are significant beyond a probability level of .001. The four pigeons therefore clearly ranked positives and negatives similarly to a degree, even though they

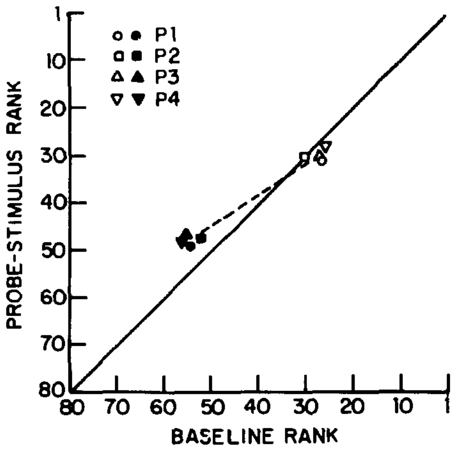


Fig. 5. Ranks assigned to probe stimuli plotted against ranks assigned to baseline stimuli, for each subject averaged across the five sessions involving probes. Open symbols for stimuli containing fish; filled symbols for those without fish. Dashed line connects the means of the open and filled symbols.

had had no prior reinforcements in their presence and were seeing them in different orders.

C. SUMMARY

Four pigeons rapidly learned to sort underwater photographs being seen for the first time according to the presence or absence of fish in them. The speed of learning and the final level of discrimination was comparable to those found in experiments using objects from a pigeon's ordinary habitat, such as trees, bodies of water, and people. Generalization of the fish category was shown by all pigeons in at least some tests, although not in every instance. Stimuli were often ranked in similar order by the pigeons, prior to any reinforcements in their presence and independent of the order in which they were seen.

IV. Experiment 3: Fish Concept versus Quasi-Concept

The primary purpose of this experiment was to compare the learning of a concept such as fish with what may be called a "quasi-concept." For a quasi-concept, a set of instances that could be sorted according to the presence or absence of fish is instead divided randomly, with as many fish in the negative class as in the positive. If the quasi-concept is harder for pigeons than the fish concept, we have further evidence that the pigeons are actually using an overarching principle of classification in learning to sort instances, when one is available.

A. METHOD

1. Subjects

Six adult, male white Carneaux pigeons at 80% of *ad lib* weight served as subjects. All the pigeons had worked in operant chambers before, but none had seen photographic slides until the first session in this experiment.

2. Apparatus

The chamber was similar to that described for Experiment 2, with one major exception. In the present experiment, the response was not to the screen itself, but to a standard pigeon key mounted to the left of the screen at a distance of .051 m. The key was operated by pecks exceeding a force of .15 N.

3. *Procedure*

As in Experiment 2, a session consisted of the presentation of 80 slides for an average of 30 sec each. Half the slides were positives and half negatives, as defined below. There was a 5-sec intertrial interval between slides during which the screen and key were dark. The response key was illuminated with white light 5 sec after the slide was projected on the screen. A VI 30 sec programmed reinforcements in the presence of positives, and reinforcements were 2.5 sec of food hopper time. There was a penalty for pecking in the presence of negatives: a negative was terminated no sooner than 10 sec after a peck. Therefore, by pecking in the presence of a negative, a pigeon may have been extending the time during which it could earn no reinforcements. As in Experiment 2, the rates of pecking used for analysis were based only on responding up to the first reinforcement in the presence of a positive and, for a negative, up to where the first reinforcement would have been if it had been positive instead of negative.

Starting with the first session, all pigeons saw the same 80 slides daily, randomized anew for each subject and each session. The 80 slides were, in fact, the slides used as generalization probes in Experiment 2 and for the final generalization test plotted as session 18 in Fig. 4. These stimuli, it should be recalled, were among those used in Experiment 1, and they represent the full range of human reaction times. For three of the pigeons, the positives were the 40 stimuli containing fish and the negatives were the other 40. For the other three pigeons, the 80 stimuli were randomly divided into a positive and a negative set, with the proviso that each set should include 20 fish stimuli. All pigeons in this quasi-concept group had the same set of positives and negatives. Once a stimulus was classified as positive or negative, it remained so for the entire experiment.

The training phase for each subject lasted until it had five consecutive sessions in which the index of discrimination was statistically significant at the .05 level or better. Then, there was one session with 80 new stimuli (40 fish and 40 non-fish) drawn at random from the collection of 240 used for training in Experiment 2. For the fish-concept group of pigeons, the presence of fish was the signal for the reinforcement schedule. For the quasi-concept group, the 40 fish and 40 non-fish slides were randomly assigned as positive or negative with the constraint that 20 fish and 20 non-fish slides end up as positives. The generalization session followed the same procedure as in training with two exceptions: 10 sec of any positive stimulus had to elapse before a reinforcement could be delivered, and there was no penalty for pecking at the negative slides.

B. RESULTS AND DISCUSSION

The fish-concept group should presumably show the benefits of the concept by learning to sort stimuli quicker than the quasi-concept group. This is, in fact, apparent from the descriptive measures in Table II. Although all the pigeons were seeing the same stimuli, the ones in the fish-concept group significantly discriminated earlier than the ones in the quasi-concept group by either of two plausible measures. The first significant discrimination (.05 level) for the fish-concept group was on the 2.7 session, averaging over pigeons, versus on the 13.3 session for the quasi-concept group. The last nonsignificant session prior to meeting the criterion of five significant sessions was 11 and 24.7, averaged for the fish-concept and quasi-concept groups, respectively.

Subjects learned to sort fish more rapidly than to sort into two arbitrary classes, but there was no substantial difference in the level of discrimination reached during the five criterial sessions. The average value of r_p for the fish-concept group was 32.2, negligibly higher than the average of 31.7 for the four pigeons on session 18 in Experiment 2, which is when they saw these same stimuli. Seeing the stimuli daily, as was the case for the pigeons in the present experiment, conferred no obvious advantage over the pigeons in Experiment 2 who were seeing the stimuli for only the second time and had had no reinforcements in their presence the first time they were seen. The average value of r_p for the quasi-concept pigeons on the criterial sessions was 32.7. Though they learned more slowly they achieved roughly the same level of discrimination as the fish-concept pigeons.

Since the subjects were seeing the same stimuli daily, since the positive

TABLE II

SESSIONS TO ACQUISITION OF FISH CONCEPT AND QUASI-CONCEPT

	Fish concept subjects			Quasi-concept subjects		
	P4	P5	P6	P7	P8	P9
Session of first significant[a] discrimination	2	4	2	14	15	11
Session until last nonsignificant discrimination	14	5	14	25	28	21

[a] $p < .05$.

stimuli were randomized with respect to the reinforcement schedule, and since all stimuli were randomized with respect to order and duration of presentation, it becomes relevant to determine whether stimuli were consistently ranked high or low. If positives and negatives had equal (or equivalent) treatment within the experiment, then a consistent pattern of ranking would suggest a role for the properties of the pictures as such (rather than, say, their reinforcement histories).

The first test of this possibility is presented in Table III, consisting of the concordances for each pigeon in both groups for the five criterial sessions. For this purpose, positives and negatives were ranked separately for the five sessions and Kendall's *W* provides a convenient test of a pigeon's session-by-session consistency within each category. Only P6 failed to reach statistically significant concordance; the other five subjects were significantly concordant (at $p < .01$) to about the same degree for both positives and negatives. It may be relevant that P6's discrimination at criterion was the poorest among all subjects, averaging an r_D of 35.2.

Table III suggests either that stimuli vary in inherent discriminability or that the accidents of each pigeon's early experimental history with the stimuli produce an ordering of discriminability that lasts into the criterial sessions. Table IV, in effect, excludes the latter interpretation by the significant concordances across subjects in each group. For each subject, an average order of stimuli was determined by ranking the sum of the ranks over the criterial sessions, for positives and negatives separately. These ranks were used for calculating the concordances in Table IV. The pigeons, it should be recalled, each saw the stimuli in a different order, with different presentation times and number of reinforcements for the positive stimuli. The agreement among the ranks in each group arises, it would seem, from inherent properties of the stimuli themselves.

Further evidence to this effect is in the comparison of the rankings of the pigeons in Experiment 2 on session 18 with those of the fish-concept

TABLE III

WITHIN-SUBJECT CONCORDANCE (KENDALL'S *W*)

	Fish concept subjects			Quasi-concept subjects		
	P4	P5	P6	P7	P8	P9
Positives	.422**	.406**	.215*	.380**	.403**	.350**
Negatives	.645**	.396**	.250*	.457**	.318**	.345**

*n.s. ($p > .05$).
**.01 > p.

TABLE IV

BETWEEN-SUBJECT CONCORDANCE (KENDALL'S W)

	Fish concept	Quasi-concept
Positives	.500*	.539**
Negatives	.615**	.555**

$^{*}.01 < p < .05$.
$^{**}p < .01$.

group here. The same slides were involved, but the experimental histories were different, since the pigeons in Experiment 2 had been trained on daily samplings from a pool of 240 slides that included none of those used here. The Spearman rank–order correlation between the ranked sum of ranks for the four pigeons in Experiment 2 and the three pigeons in the fish concept group was .466 for positives and .667 for negatives, both significant beyond the .01 level. Since these correlations are based on within-category rankings, they do not include the effect of the discrimination between positives and negatives as such. Rather, they indicate that a pigeon's difficulty in sorting positive or negative instances of "fish," as operationalized in the measures we use, is independent of whether the instances had been seen before.

Tables III and IV indicate within-category concordance for the quasi-concept procedure, which seems harder to interpret than the concordance among subjects sorting between fish and non-fish. What can be the basis of concordance with 40 randomly selected underwater photographs, of which 20 happen to include a heterogeneous assortment of fish? Mostly, we must leave this question unanswered, but Table V may contain a clue. This table presents correlations for rankings of the 80 slides assigned to four equal subsets: those containing fish versus those not containing fish, each subdivided into whether they were positive or negative for the quasi-concept procedure. The correlations (Spearman) are for the subjects in the fish-concept procedure versus those in the quasi-concept procedure, of which only one is significant. Rankings of fish as positives were negatively correlated in the two procedures. If the fish-concept subjects assigned high ranks to "good" exemplars of fish, then the quasi-concept subjects must have tended to assign them low ranks. This would be expected if the good exemplars of fish as positives were maximally difficult for the quasi-concept subjects since they most resembled the good exemplars of fish as negatives.

Table VI spells this hypothesis out more fully in idealized form. For

TABLE V

SPEARMAN RANK-ORDER CORRELATIONS FOR FOUR SUBSETS OF THE 80 STIMULI

	Slide content	
	Fish	Non-fish
Contingency of reinforcement		
Positives	−.457**	.223*
Negatives	.210*	−.207*

*n.s. ($p > .05$).
**$p < .05$.

each subset of stimuli, it shows the predicted ranking of good to bad exemplars of fish and non-fish for each procedure, and it also shows the location of the ranks with respect to the mid-rank of the whole session. To make the hypothesis clearer, it also assumes, counterfactually, that the subjects in both procedures discriminated perfectly, ranking all positives above, and all negatives below, the mid-rank. Let us now suppose that all subjects noticed resemblances among at least some fish stimuli and also

TABLE VI

PREDICTED CORRELATIONS ACROSS THE TWO PROCEDURES

	Fish positive[a]		Fish negative[a]		Non-fish positive[a]		Non-fish negative[a]	
	Fish concept	Quasi-concept	Fish concept	Quasi-concept	Fish concept	Quasi-concept	Fish concept	Quasi-concept
Ranks								
	Good	Bad	Good			Bad		
	Bad	Good	Bad			Good		
Mid-rank								
				Good	Bad		Bad	Good
				Bad	Good		Good	Bad
Correlation								
Predicted	−		+		+		−	
Observed	−.457		.210		.223		−.207	

[a] Heading gives the reinforcement contingency for the quasi-concept procedure.

resemblances among at least some non-fish stimuli. The fish-concept group would benefit from any resemblances noted, while the quasi-concept group would be misled by them. If good exemplars are those that resemble the largest variety of other instances, then it follows that good exemplars of fish and non-fish should be further from the mid-rank than bad exemplars for the fish-concept procedure and vice versa for the quasi-concept procedure, as indicated in Table VI.

The predicted directions of the correlation are shown at the bottom of the table, along with the observed values. Although only one correlation was statistically significant, all four were in the predicted direction. The observed correlations would be attenuated by the extent to which discrimination was imperfect, which was, as noted above, considerable. It would be more appropriate to test this hypothesis with a concept or a procedure that yields higher values of ρ than found here.

We will, in fact, conclude the description of Experiment 3 by considering the possible reasons why the level of discrimination might have been as low as it was here. First, we present a further bit of evidence that it was lower than typical. On the session following the criterion sessions, each pigeon was shown 80 new slides (40 fish and 40 non-fish) drawn at random from the pool of 240 slides used for training in Experiment 2. All three fish-concept subjects showed a significant discrimination between fish and non-fish (average $r_p = 34.6$) and none of the quasi-concept subjects did (average $r_p = 41.1$). The significant discriminations in the fish-concept group indicate some sort of generalized rule (or rules) for classification, but a value of 34.6 corresponds only to $\rho = .65$, which is low even by the standard of Experiment 2.

Discrimination may have been relatively poor for the fish-concept group for a variety of reasons, besides any inherent difficulty in the sample of slides used to train them (a possibility already considered in Experiment 2). First, the subjects saw the same instances in every session. Presumably, for some concepts, if not all (see Cerella, 1979, for a counterinstance), the subject must be exposed to variable exemplars of a class before it induces a class rather than a description of the individual alone. Between the unique individual and a generalized class may come intermediate levels of categorization that may produce poorer discrimination than the extremes. It may, for example, be easier to recognize a particular Italian man or men in general than Italian men in general, although with enough practice and a sufficiently motivating contingency, it is probably even possible to distinguish native Florentines from Neopolitans at some minimal level of accuracy. An individual exemplar is nested within an indefinitely large number of superordinate categories, only a small fraction of which we can name or otherwise differentiate at a given point in training. It is therefore not a priori unlikely that the same 40 heterogeneous examples of fish seen repeatedly may be at an awkward intermediate level for inducing a category that covers fish in general.

Besides that possible reason for the relatively poor discrimination several others should be noted. Inexperienced subjects were used in Experiment 3. The pigeons in Experiment 2 had previously learned to sort for trees, which may have left some residual skill useful in sorting for fish. As yet, there appears to be no published evidence for such a perceptual skill in animal subjects, but it has been reported (Deregowski, Muldrow, & Muldrow, 1972) that inexperienced human beings must first learn to make sense of photographs. On the other hand, inexperienced pigeons learned (Herrnstein, 1979) to sort for trees rapidly and well. If general picture-perceiving skill is involved, it apparently interacts with something about the category being learned, perhaps its difficulty. Next, the pigeons pecked the stimulus screen itself in Experiment 2, but an adjacent key in Experiment 3. It has been suggested (Ferster, 1964) that stimuli on an operandum gain better control than stimuli elsewhere, perhaps because of the closer contiguity between stimulus and response. Also, the food hopper cycle in Experiment 2 lasted 5 sec and only 2.5 sec in Experiment 3, and it is at least conceivable that this might affect discrimination level. Finally, there was a 10-sec delay of termination for pecking during negatives in Experiment 3 but no penalty in Experiment 2. So little is known about the sources of variation in learning open-ended concepts that we cannot exclude this minor difference between the two experiments.

C. SUMMARY

Pigeons learned to sort 80 stimuli into two categories. They learned quicker when the categories corresponded to fish versus non-fish than when the same stimuli were divided into two arbitrary categories. Within each procedure, there was concordance within and between subjects for each category. The subjects in the fish-concept procedure were concordant with pigeons in Experiment 2, who saw the stimuli as a generalization test after being trained with other underwater photographs.

V. General Discussion

Seven out of seven pigeons in two experiments learned to sort underwater photographs on the basis of the presence or absence of fish. All the pigeons generalized significantly to new sets of underwater photographs, at least on some tests. The pigeons agreed to a degree on the ranking of the stimuli within the category of fish and the category of non-fish, and did so even when they were trained on different sets of exemplars, though all the pictures were taken in similar natural habitats. An additional group of three pigeons learned, but more slowly, a "quasi-concept," an arbitrary division of a set of 80 stimuli. Although these subjects were concordant with each other, they disagreed in rankings with the pigeons in the fish-concept group. However, the disagreements themselves were shown to be consistent with a fish-category of some sort even in the quasi-concept group (see Table VI).

A. COMPARISON WITH HUMAN SUBJECTS

In light of these results, it would be hard to deny pigeons a concept of sufficient generality to correspond at least in part with the human concept called fish. But this conclusion stops far short of saying that the pigeons' concept is equivalent to the humans'. Table VII shows that they are, in fact, not even significantly correlated by our measures. The composite rankings of the four pigeons in Experiment 2 and the three pigeons in the fish-concept group in Experiment 3 were compared with that of the 12 human subjects in Experiment 1, for the 80 stimuli (40 fish and 40 non-fish) common to the three experiments. Although there was substantial agreement within species, none of the correlations between species reached statistical significance.

After examining the 80 photographs, we believe there are two main reasons that pigeon and human rankings failed to correlate. First, people quickly infer fish from clear, partial views, but pigeons do not. The partial fish in the slide reproduced as Fig. 6 earned a composite rank of 6.5 out of 40 for the humans and 38 and 28 for the pigeons in Experiments 2 and 3. Figure 6 was thus a good exemplar for the human subjects and a poor to middling exemplar for the pigeons. Second, the pigeons had relatively less trouble than the people with atypical fish, as long as they were represented in the sample seen in training. Humans take a while to

TABLE VII

SPEARMAN'S RANK-ORDER CORRELATIONS BETWEEN HUMAN AND PIGEON RANKINGS

	Experiment 1 vs 2	Experiment 1 vs 3
Fish	.045[a]	−.128
Non-fish	.243	.181

[a] All correlations not significant at $p \leq .05$.

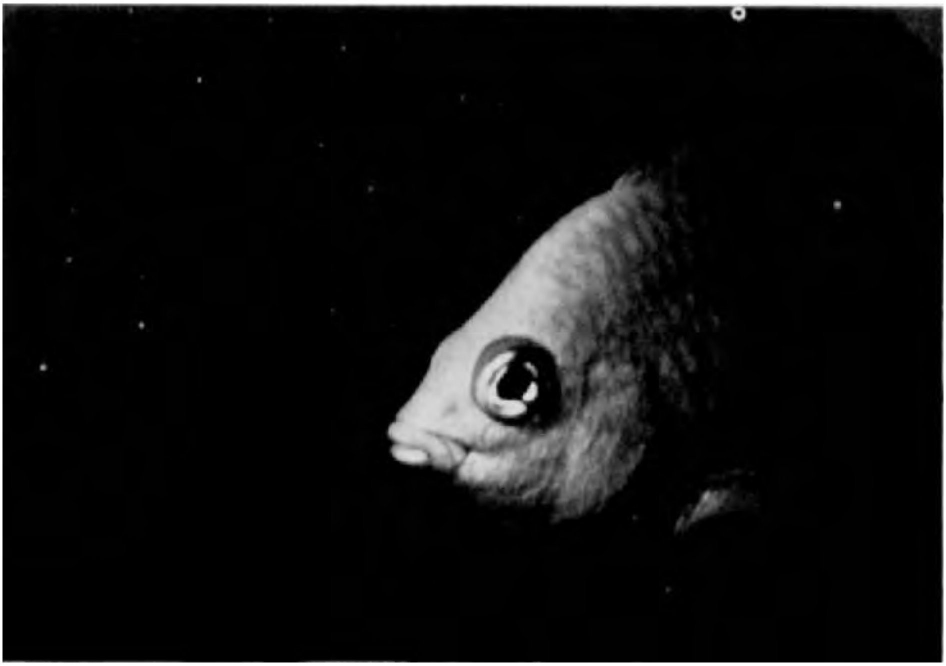


Fig. 6. A good fish exemplar for the human subjects and a middling to poor one for the pigeons. Figures 6–9 originally in color.

decide that a skate (see Fig. 7) counts as a fish, but pigeons, who may have few if any preconceptions about prototypical fish, consider Fig. 7 a good positive instance. Figure 7 was ranked thirty-first out of 40 by the human subjects, but fourth and sixth by the pigeons in Experiments 2 and 3. Several other discrepancies in rankings by pigeons and people seemed to be of this type.

Figures 8 and 9 show the three best and worst fish and non-fish for Experiments 2 and 3, based on the composite rankings across both experiments. Perhaps the most remarkable result is in Fig. 8, for it shows that the highest pigeon ranks are for stimuli presenting a canonical view of a fish by human standards. We might have attributed the human canonical view to familiarity with side views in illustrations, on fish counters, etc., but side views of fish were not especially abundant in our collections of photographs, many of which contained fish only incidentally. Side views may be canonical because they are easy rather than easy because they are canonical. The side view presents a recognizable and memorable form, and that may be the reason it becomes common in books, drawings, etc. Figures 8 and 9 also show that the pigeons were influenced by figure-

Fig. 7. A poor fish exemplar for the human subjects and a good one for the pigeons.

ground relations, as we noted people were. Busy slides containing ambiguous forms that may or may not have been fish were troublesome for people as for pigeons. The pigeons tended to mistake other animals for fish more often than the humans, but, more often than not, they classified turtles and divers (see Fig. 9) correctly.

Three factors were mentioned earlier as contributing to reaction time for the human subjects—contrast or figure-ground, prototypicality, and canonical view. From the examples in Figs. 6–9, we can see that the first and third also contribute to the pigeon's performance, but that the second, prototypicality, partially differentiates pigeons and people. The human ability to infer a whole fish from a partial view further differentiates them. The insignificant correlations in Table VII should not, then, be interpreted as a total discrepancy between the pigeon's concept and the people's, but as the result of specific differences. Although we have not tried it, we believe we could pick a sample of stimuli that would produce substantial correlations between human and pigeon rankings, by eliminating nonprototypical fish and partially obstructed views.

An obvious explanation of the difference in prototypicality is that people have prior experience of fish and pigeons do not. Another is that people bring biological knowledge to bear on whether an instance is a fish or not, and would, for example, reject whales while pigeons probably would include them. People also include whales among the fish until they are taught to resist the spontaneous tendency to do so. The pigeon's category of fish is undiluted by preexperimental exemplars and also by formal knowledge. For the pigeons the category seemed to be anchored by the positive and negative stimuli and then extended according to ordinary perceptual dynamics. This, in turn, implies that the pigeon's category was not significantly shaped by genetically programmed exemplars

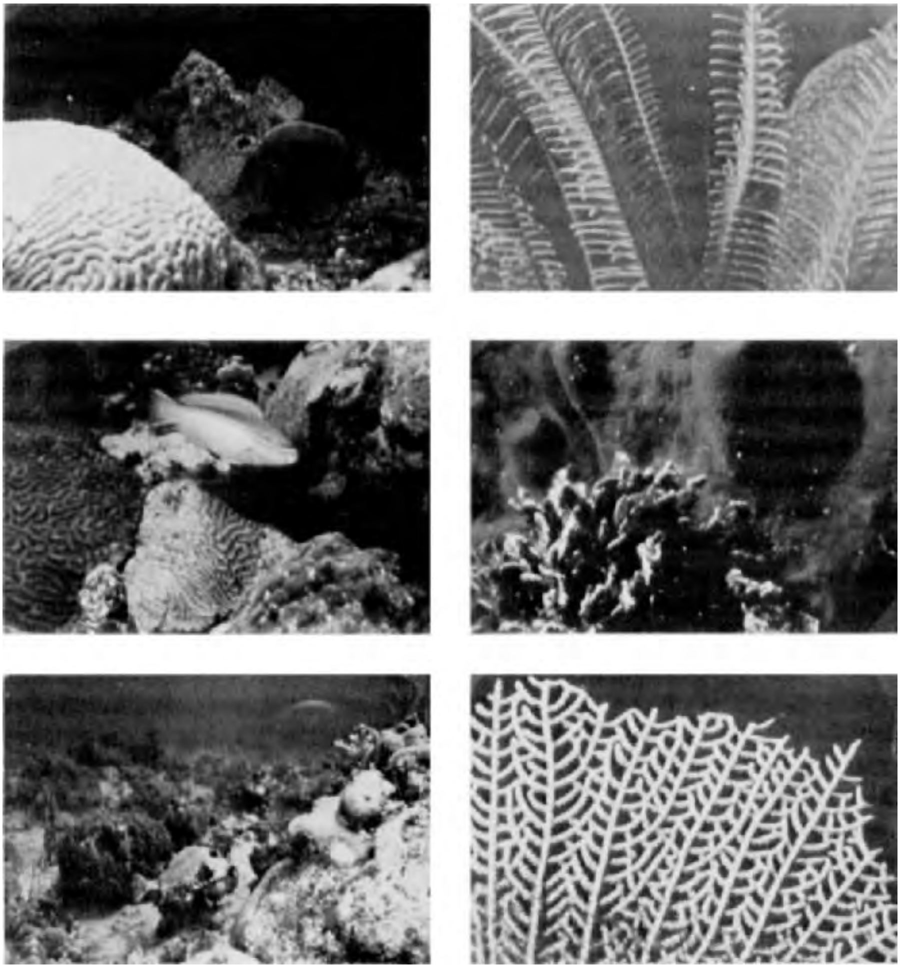


Fig. 8. Based on pooled rankings by the pigeons in Experiments 2 and 3, the three best fish (left) and non-fish (right) exemplars.

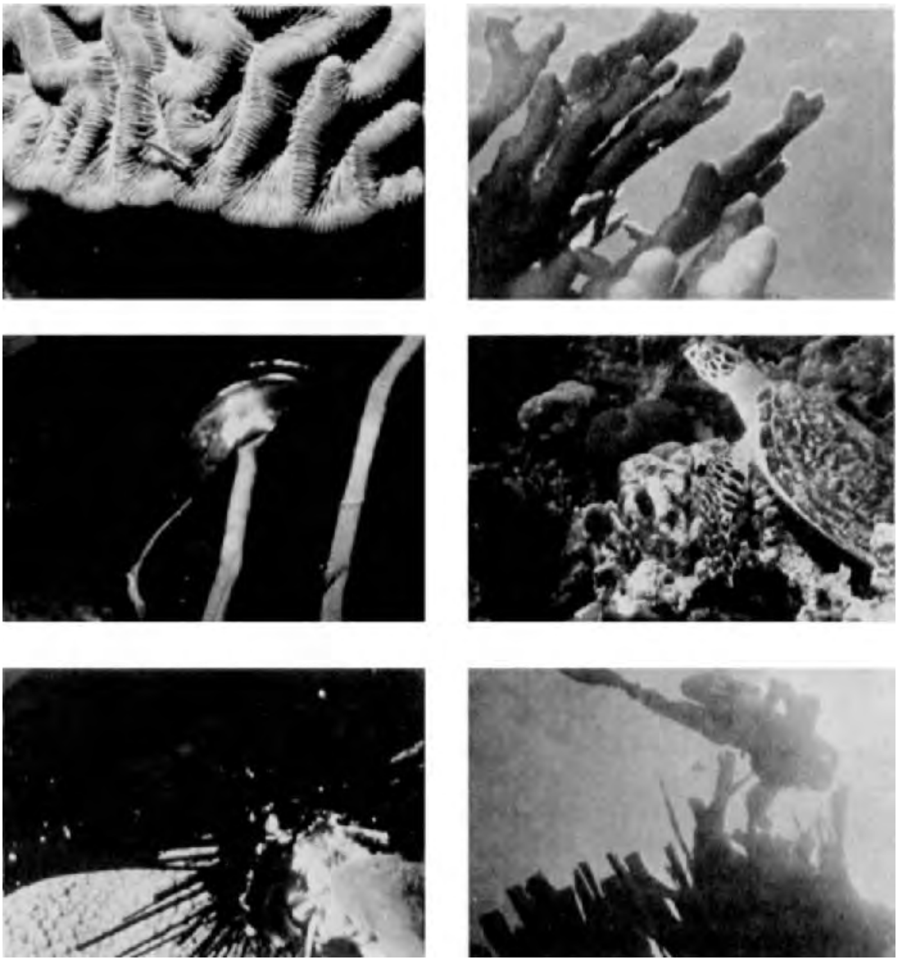


Fig. 9. Based on pooled rankings by the pigeons in Experiments 2 and 3, the three worst fish (left) and non-fish (right) exemplars.

unless the photographs in the sample happened, by some infinitesimal chance, to correspond to them.

B. PERCEPTUAL FACTORS

Rather than providing prototypes, the genetic constraints are more likely to be expressed in the general perceptual factors affecting categorization, of which figure-ground and canonical view are examples already noted. Pigeons and their precursors have needed to be able to recognize animals at approximately taxonomic levels of classification, conspecifics as well as nonconspecifics. Concretely, this is a problem of dealing with variability, for the members of taxonomic classes may be far from identical, yet still be functionally equivalent from the standpoint of a responding organism. Mice, for example, ought to flee from any cat, since cats will pounce on any mouse. The contingencies of reinforcement, themselves, tend to shape the adaptive complex categories, but it would be surprising if animals' perceptual systems had not evolved with the innate flexibility to categorize instances as variable as fish and other classes of animate and inanimate objects. The quasi-concept procedure illustrates that some subsets of the natural environment are easier to categorize than others, without regard to differences in reinforcement contingencies. Something in the pigeon's perceptual dynamics ties fish together as a class, prior to differential reinforcement. It is immaterial that pigeons do not encounter fish outside the laboratory. We would expect the perceptual dynamics to equip the pigeon more for possible environments than for the much narrower class of actual environments, just as its appetite is equipped for possible diets rather than for specific foodstuffs. Fish may not seem too probable an object for pigeons, but then the discriminations here were not as good as they have been in the experiments using trees, oak leaves, and human beings as the stimulus classes.

Whatever the underlying perceptual laws are, they are not necessarily making a single class out of the positive exemplars. From an assortment of fish pictures, the pigeons may be inducing multiple positive types, as far as we can tell from the data. The union of several distinct types of fish forms could handle the contingencies of reinforcement as well as a single class accommodating most all fish, and it may be a more feasible task for a perceptual system. Recent studies with human adults suggest that in many concept formation situations humans also form heterogeneous categories on the basis of sets of exemplars rather than a single class defined by criterial attributes (Brooks, 1978; Medin & Schaffer, 1978). Detailed analyses of acquisition and generalization, which we have not done here, could provide clues about a subject's principles of classification, but the point may be worth acknowledging in the abstract. When a perceptual system interacts with a set of varying exemplars, the resulting category structure depends on the match between the properties of the system and the exemplars.

A single example may help. For human beings, the category "mammal" is perceptually diverse, a good example of the union of multiple types stretching from tiny voles to giant whales. In contrast, according to von Uexküll (1957), blood-sucking ticks identify mammals on the basis of the presence of butyric acid, a volatile compound found in all mammalian flesh. What is diverse for us is integrated by the character of the tick's perceptual adaptation to its favorite habitat.

We would suppose that the variety of fish in the training pictures probably resulted in multiple types for the pigeons, good enough to yield significant discrimination and generalization, but not good enough to accommodate every other fish in the collection. A pigeon may not recognize an eel as a fish if it had never seen eels before and had seen snakes and sea-cucumbers as negative instances. However, a few eels in training seems to have been enough here to add eels to the list of positive types, which then generalized so that the one sea-cucumber in the set was usually misclassified as a fish by all the pigeons that saw it. With further nonreinforcement for pecking at sea-cucumbers, the pigeons could probably learn to distinguish them from eels.

The reinforcement contingency may tell the pigeon how to sort the objects it sees in the stimuli, but not what objects it sees. The objects it sees seem to be determined by the interaction of its perceptual system with the exemplars, rather than by experimental procedures. This was also the conclusion of earlier studies (Herrnstein, 1979; Herrnstein & Loveland, 1964) in which pigeons formed categories for people and trees. But unlike people and trees, fish do not inhabit the pigeon's natural environment. Consequently, the present conclusion is not merely a replication. It further implies that the limited information in a picture is enough for a pigeon to form categories for objects found neither in its, nor its recent ancestors', environment. The present results add evidence of the sheer power of the pigeons' capacity for categorization, and it is likely that the pigeon is not exceptional in this respect.

C. NATURAL CATEGORIES ARE EASY

In most of the animal research on open-ended categories, the discrimination forms rapidly. Pigeons, for example, come to discriminate trees from non-trees or fish from non-fish more rapidly than they do among pure tones or monochromatic lights (Herrnstein, 1979). The physical simplicity of tones or patches of color may tax a perceptual system geared to the level of natural functioning, which is a level of three-dimensional objects, complex sounds, and multisensory happenings. A photograph of a tree or an animal of some type or a body of water comes closer to the congenial level of classification than abstract entities such as pure tones or patches of color. However, trees and the like, though three-dimensional, are being shown in two-dimensional view. The pigeons apparently quickly infer the third dimension, for if they did not, they would need to learn to correct for perspective projections within the experiment. For example, after seeing trees from one set of angles, they rapidly, perhaps immediately, generalize to trees at new angles. The variations due to perspective are transparent to human observers because they see the picture as a three-dimensional scene, and apparently they are transparent to pigeons as well. We could equally well say that people and pigeons display object constancy when looking at photographs of trees.

Cerella (1977) demonstrated that pigeons sometimes fail to infer a third dimension when it would help to do so. Trained on a limited range of views of a cube, pigeons failed to generalize to new views. To human observers, it is so obvious that the new and old views are of the same object that it would be quite hard to identify which stimuli are new and which old. This is another example of transparent variation, but only for people. For pigeons, the new views are likely to be treated as new stimuli, not members of the category of cubes.

In Cerella's cube experiment, the stimuli were copies of computer-generated line figures. They were not scenes from a three-dimensional space, even though the perspective relations among the lines mimicked the edges of a real cube rotated in space. Nevertheless, the stimuli failed to trigger the pigeon's proven capacity for object-constancy, so it apparently saw the stimuli as just so many line segments on a plane, with no unifying invariance. It would be interesting to know if the pigeons could learn the category of cube from photographs of, say, an actual cubical block at different angles. If so, we could conclude that the pigeon needs more context than we do before it can take advantage of its innate (or otherwise preexisting) knowledge of perspective. But even human subjects need a certain amount of context. Not everyone, for example, can see the three-dimensional relationships in architectural drawings. A line figure of a cube for a pigeon may be like an architect's complex schematic for us.

D. WHAT IS A NATURAL CATEGORY?

Among the relatively hard problems for animals seem to be the nonnatural categories, whether they involve the simple tones, lights, and geometrical figures of traditional research on discrimination learning or Cerella's cube experiment and his Charlie Brown experiment (1975). Charlie Brown, the *Peanuts* cartoon character, proved to be a hard category for pigeons to learn. Yet, pigeons learn without special difficulty to recognize photographs containing a particular woman (Herrnstein *et al.*, 1976) in many different settings and wearing different clothing. A real person in real environments is a more demanding challenge to an artificial intelligence device (e.g., Winston, 1977) than Charlie Brown, but vice versa for pigeons.

The difference presumably has something to do with the naturalism of a real person in ordinary settings, but naturalism is not a well-defined theoretical construct. Pigeons could distinguish white oak leaves from other leaves (Cerella, 1979) even though they probably had had little if any prior experience with the leaves of any trees, let alone white oaks. Few pigeons have had prior occasion for recognizing individual people, although laboratory pigeons may come to know their caretakers. As we noted earlier, neither pigeons nor their recent ancestors have often encountered fish. In short, while fish and the like may be naturalistic, the relevant sense is that they are categories formed (i.e., by the human experimenters) via interactions with a natural environment. But they are not naturalistic in the sense that the categories formed in the pigeon's own interactions with its environment prior to becoming a subject in our experiment nor are they necessarily categories used by pigeons generally in their natural environment.

But that distinction is not yet the whole story. Pigeons who learn to recognize oak leaves would probably be fooled by the "leaf" in Fig. 10, which we made up after looking at true exemplars like those in Fig. 1. There may be no oak leaf just like our creation, but it looks like one anyway to us, and probably would to a pigeon as well. Naturalistic

Fig. 10. Fictitious white oak leaf.

categories are not entirely populated by real objects, since they also include fictional objects that satisfy the requirements for membership, whatever they are. Moreover, if an organism can form a category for horses, it probably also can for unicorns. Unicorns are naturalistic in the sense that if they existed, they would probably be picked out and classified as such, close to horses but clearly something different. The naturalism of categories does not mean that animals know ahead of time what kinds of objects the environment actually contains. Evolution prepares us just as well for a world containing unicorns and no horses as vice versa.

Inversely, there are countless uncongenial ways to divide up the natural environment. It might be difficult for us, for example, to classify objects so that their surface areas plus their average reflectances added to a constant. Difficult for us, that is. For some other creature, this invariance could make sense while our invariances of form might be incomprehensible or ridiculous. To the hypothetical creature, its invariant would seem naturalistic, even if it could not formulate the rule explicitly. Though categories may be naturalistic, they are also arbitrary.

Our naturalistic categories are a subset of a much larger set of partitionings of the sensory inputs from the environment. Human constructions, like Charlie Brown and computer-generated cubes, are harder categories for pigeons than people. After looking at some of the extremely difficult slides of moths used by Pietrewicz and Kamil (1977), we would guess that human subjects could not categorize moths as accurately as the blue jays in their study, at least at first. Species differences would be expected on evolutionary grounds, but so are the commonalities among species. The mottled coat that hides a Dalmatian dog against a natural background from its prey also hides it from us. This adds some anecdotal evidence to the experimental findings that show some commonality in the categorizations of different species. When different species inhabit distinct but overlapping environments, it is not surprising that their categorizing tendencies are also distinct but overlapping. But whether the tendency is specific or general, it consists of making certain variations in a set of instances transparent and others highly significant. Naturalistic categories may be defined as categories whose members vary transparently. For this to happen, there must be a correspondence between the variations among the members and the variations overlooked or forgotten by an observer.

E. GENERALIZATIONS IN QUALITY SPACES

When instances of a category vary only slightly, we do not look further for an explanation of the process of categorization. In fact, category formation would no doubt be subsumed under simple recognition if instances always approximated each other in some standard physical sense. Most telephones almost exactly mimic one or another of a small number of standard types. Recognizing a new instance of a familiar type of telephone would not be taken as evidence of a generalized telephone category. But when the physical variation among the category members is large relative to the variation across the category boundary, interesting issues arise concerning categorization.[3] Occasionally, large within-category variation cannot be easily explained by the observer's past experience. A child who identifies a stylized drawing of a telephone in her very first picture book exemplifies a central problem of categorization. In fact, apes (Davenport & Rogers, 1971) have been shown to be able to generalize from pictures to objects and to match a sample object presented to the sense of touch with a target presented visually (Davenport, Rogers, & Russell, 1975). Pigeons also can transfer a discrimination between objects to pictures of the objects (Cabe, 1976).

Herrnstein (1979) showed that when pigeons learn to categorize photographs of trees, the level of discrimination of an individual instance is uncorrelated with its own previous association with reinforcement, even including instances associated with no previous reinforcement at all. Virtually as soon as the pigeons discriminated any instances at all, there was evidence of a sufficiently general tree category to encompass new instances that varied widely, in simple physical terms, from the reinforced exemplars. Similarly, Cerella (1979) demonstrated that new instances of oak leaves (see Fig. 1) were recognized by pigeons after they had seen only one positive exemplar. Both of these findings indicate categories whose members spontaneously vary widely but appropriately. Because such results may seem to suggest innate knowledge of, say, oak leaves or trees in pigeons, they are sometimes considered counterintuitive if not simply spurious.

There is, however, a less perplexing way of thinking about these results. To a human observer, oak leaves resemble each other, as do at least some trees and fish. In fact, the three categories differ in their patterns of similarity. Trees and fish are each more like a family of forms, including varieties as different as palms, spruce, and spreading beeches or flounder, eels, and stonefish. The resemblances across varieties may range from

---

[3] The converse can also be interesting, but not directly relevant here. Sometimes, small physical differences between stimuli account for rapid transitions from one category to another. These may be characteristic of a perceptual modality, as in the sharp transitions heard in certain speech sounds (Kuhl & Miller, 1975) or in color sorting procedures, or a result of sharply differentiated contingencies of reinforcement, as in the discriminations between edible and poisonous mushrooms or between one person's face and someone else's.

substantial to nonexistent in visual appearance. For example, a subject first trained with only eels as exemplars of fish might fail to generalize to either flounder or stonefish. Its category may be closer to snaky creatures than to fish. In contrast, oak leaves are more nearly instances of a single form, and can therefore be placed at a lower level in a hierarchy of resemblances. Cerella's data (1979) suggest that pigeons need to see only one white oak leaf in order to generalize to many if not all other white oak leaves. It would be possible and perhaps useful to examine such patterns of resemblances further, but the outcome probably would not bear on the obvious point being made. Categories are formed as organisms generalize to similar stimuli, and oak leaves, trees, fish, etc. are, with the foregoing qualifications, similar.

Why, then, are findings on natural categories noteworthy, assuming they are? The answer seems to have to do with a faulty but understandable preconception about the dimensions of generalization or similarity. The typical independent variable for generalization paradigms is a standard physical variable—wave length, energy level, and angle of inclination are examples that come to mind. We expect physical proximity to be associated with psychological similarity and usually it is. Given a stimulus, b, on a physical dimension, where b is between two other stimuli, a and c, it is in fact likely that b will be psychologically more similar to (or no more different from) a and c than a is to c. This is a valid supposition about generalization gradients. However, the results with open-ended natural categories suggest that the converse is often untrue. It is not generally the case that psychologically similar stimuli are closer on a standard physical dimension than less similar stimuli. For natural categories, even relatively homogeneous ones like oak leaves, a more complex physical representation is needed than for the more familiar generalization gradients. Of the myriad possible descriptions of the shapes of leaves, we need one for which oak leaves would be closer to each other than to other leaves. Finding a satisfactory description has been like looking for a small needle in a large haystack. The preconception that is being overthrown concerns primary stimulus generalization, which proves to be inadequately accounted for by simple physical proximities.

Quine's (1969) notion of "innate quality spaces" captures the domain of primary generalization better than standard psychological usage because it presupposes little about physical dimensions. The internal representation of stimuli constitutes, for Quine, a quality space that yields generalizations, or proximities, that tend to be congruent with the functional categories of objects. The quality space is arbitrary but not accidental. It is no accident that the varying leaves produced by a species of tree happen to look similar to us. Rather, it and the many other comparable congruences between perceived form and actual function are an evolutionary adaptation not unlike the congruence between a creature's teeth and its diet. Both have been shaped by a history of past environments and both serve to delimit the individual's present environment, one perceptual and the other nutritional. Inasmuch as generalization is congruent with functional groupings, it is a reasonable approximation to innate knowledge but without the excess metaphysical overtones. It is not that pigeons know that oak trees grow a certain kind of leaf but that pigeons are at an advantage if they are interacting with an environment in which oak trees do so. The gradients are fallible; the counterfeit oak leaf in Fig. 10 would probably have been taken as valid. Pigeons would probably have mistaken whales for fish. As Quine points out, one task of individual experience is to perfect, if not entirely replace, with increasingly powerful and comprehensive categories, the implicit knowledge built into the quality space. The categories of science itself, Quine points out, eventually supplant the similarities of our quality spaces.

REFERENCES

Bamber, D. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of Mathematical Psychology,* 1975, **12,** 387-415.

Brooks, L. Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization.* Hillsdale, N. J.: Erlbaum, 1978.

Cabe, P. A. Transfer of discrimination from solid objects to pictures by pigeons: A test of theoretical models of pictorial perception. *Perception and Psychophysics,* 1976, **19,** 545-550.

Cerella, J. *Studies of pattern perception in the pigeon.* Unpublished doctoral dissertation, Harvard University, 1975.

Cerella, J. Absence of perspective processing in the pigeon. *Pattern Recognition,* 1977, **9,** 65-68.

Cerella, J. Visual classes and natural categories in the pigeon. *Journal of Experimental Psychology: Human Perception and Performance,* 1979, **5,** 68-77.

Coren, A. The unnatural history of Selborne. *Punch,* June 6, 1979.

Davenport, R. K., & Rogers, C. M. Perception of photographs by apes. *Behaviour,* 1971, **39,** 318-320.

Davenport, R. K., Rogers, C. M., & Russell, I. S. Cross-modal perception in apes: Altered visual cues and delay. *Neuropsychologia,* 1975, **13,** 229-235.

Deregowski, J. B., Muldrow, E. S., & Muldrow, W. F. Pictorial recognition in a remote Ethiopian population. *Perception,* 1972, **1,** 417-425.

Ferster, C. B. Arithmetic behavior in chimpanzees. *Scientific American,* 1964, **210,** 98-106.

Herrnstein, R. J. Acquisition, generalization, and discrimination reversal of a natural concept. *Journal of Experimental Psychology: Animal Behavior Processes,* 1979, **5,** 116-129.

Herrnstein, R. J., & Loveland, D. H. Complex visual concepts in the pigeon. *Science,* 1964, **146,** 549-551.

Herrnstein, R. J., Loveland, D. H., & Cable, C. Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes,* 1976, **2,** 285-311.

Kuhl, P. K., & Miller, J. D. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science,* 1975, **190,** 69-72.

Medin, D. L., & Schaffer, M. M. Context theory of classification learning. *Psychological Review,* 1978, **85,** 207-238.

Pietrewicz, A. T., & Kamil, A. C. Visual detection of cryptic prey by bluejays (*Cyanocitta cristata*). *Science,* 1977, **195,** 580-582.

Quine, W.V.O. Natural kinds. In W.V.O. Quine (Ed.), *Ontological relativity and other essays.* New York: Columbia University Press, 1969.

Rosch, E. Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization.* Hillsdale, N. J.: Erlbaum, 1978.

von Uexküll, J. A. stroll through the worlds of animals and men: A picture book of invisible worlds. In C. H. Schiller (Ed. and trans.), *Instinctive behavior.* New York: International Universities Press, 1957 (German original, 1934).

Winston, P. H. *Artificial intelligence.* Reading, Mass.: Addison-Wesley, 1977.

This Page Intentionally Left Blank

FREEDOM OF CHOICE: A BEHAVIORAL ANALYSIS[1]

A. Charles Catania

THE UNIVERSITY OF MARYLAND BALTIMORE COUNTY
CATONSVILLE, MARYLAND

I. Freedom of Choice

> In this then consists Freedom, (*viz*) in our being able to act, or not to act, according as we shall choose, or will.
>
> John Locke

---

[1] Research supported by NIH Grant MH-18506 to New York University, and by NSF Grants GB-43251 and BNS76-09723 to the University of Maryland Baltimore County. The students who assisted in the research at both campuses are too numerous to list, but the work would not have been accomplished without them. Of the many colleagues with whom this research was discussed, special mention is due Thom Verhave, whose speculations about the effectiveness of the schedules described in Findley (1962) prompted the initial experiments. Some of these data were presented in an invited address (W. N. Schoenfeld, chair) at the American Bicentennial meeting of the Eastern Psychological Association (New York, April 1976).

ISBN 0-12-543314-X

What is freedom? Freedom is the right to choose.

ARCHIBALD MACLEISH

Freedom from something is not enough. It must also be freedom for something. Freedom is not safety but opportunity.

ZECHARIAH CHAFEE, JR.

Like many another abstract term, *freedom* has been variously defined. It has been equated with free will, with the absence of restraint, with liberty or license, and with the availability of alternatives. Dictionaries have included the capacity for choice among its properties. In philosophy, the relation between freedom and determinism has been a recurring issue (e.g., Berofsky, 1966; Enteman, 1967). For example, the compatibility of freedom and determinism has been argued on the grounds that free and responsible decisions are impossible within a context of indeterminate events (Greenspan, 1978; Hobart, 1934). Implicit in such arguments is the question of whether the determinants of behavior are to be found inside or outside of the individual. Perhaps for this reason, the freedom–determinism issue became interlocked with that of freedom from coercion as the philosophical treatment of freedom evolved in such writings as those of Locke (1690), Hume (1748), Mill (1859), and Skinner (1971). Thus, the political implications of different views of freedom began to figure more and more prominently in philosophical debate.

In his controversial *Beyond Freedom and Dignity,* Skinner (1971) considered both the freedom-determinism issue and freedom from coercion, and discussed how the literature of freedom has served to counter certain types of behavioral control. With the argument that behavior can be manipulated or controlled even in situations in which it would not ordinarily be regarded as coerced, Skinner identified some of the limitations of traditional concepts of freedom. But whether behavior is regarded as free or controlled or coerced, it involves choice, and choice implies the availability of alternatives. We may therefore interpret questions about the value of freedom as questions about preference for the availability of alternatives (e.g., Catania, 1975; Partridge, 1967; Reese, 1966, pp. 62-63; Voss & Homzie, 1970). Such questions are empirical: Are free choices preferred to forced choices? If so, what are the behavioral consequences of these preferences? Are they products of ontogeny or phylogeny? Are they limited to humans, or can they be demonstrated with other organisms?

The present treatment addresses these questions in a sequence of experiments with pigeons. With these nonhuman organisms, we need be less concerned that any preferences we might observe are simply products of our particular human culture. Our first objective is to demonstrate free-choice preference. Once the preference has been demonstrated, our next step is to explore some of its properties and to identify some of its limiting conditions. These limiting conditions may help us to refine our definition of choice. We may then be able to return to the concept of freedom better equipped to consider its character and its implications.

II. Demonstrating Free-Choice Preference

> . . . "free will" is the awareness of alternative choices.
>
> ARTHUR KOESTLER

> Man's struggle for freedom is not due to a will to be free, but to certain behavioral processes characteristic of the human organism. . . . The literature of freedom . . . has made the mistake of defining freedom in terms of states of mind or feelings.
>
> B. F. SKINNER

In an experiment on the value of choice, Voss and Homzie (1970) measured food-deprived rats' preferences for two paths leading to a sucrose solution. One path consisted of a single route; the other allowed the rat to choose between two subpaths. A preference for the path with a choice of subpaths was shown by 14 of 15 rats. The experiment did not separate the availability of alternatives from the different stimulus properties of routes with or without subpaths, nor did it exclude the possibility of preference for exploration or for variability (e.g., Dashiell, 1925; Krechevsky, 1937). Nevertheless, the findings were consistent with the interpretation that organisms prefer the availability of a choice. The present experiments examined the preference for alternatives in pigeons. Pecks on one of two keys occasionally produced a single route to a food reinforcer; pecks on the other key occasionally produced a choice between two routes to the same food reinforcer.

A. THE CONCURRENT-CHAIN SCHEDULES

In the initial experiments (Catania, 1975), three adult, male, White Carneaux pigeons were maintained at about 80% of free-feeding weight and served in daily sessions in the apparatus illustrated in Fig. 1. Of the six translucent pigeon keys, the two lower keys could be lit white, the two upper left keys blue or amber, and the two upper right keys green or red, by 6-W Christmas-tree lamps mounted behind them. Each key operated at a minimum force of about .14 N. Pecks on lit keys produced feedback clicks; pecks on dark keys had no scheduled consequences. During reinforcement, the 3-sec delivery of grain in the standard feeder, the feeder was lit and all key lights were off.

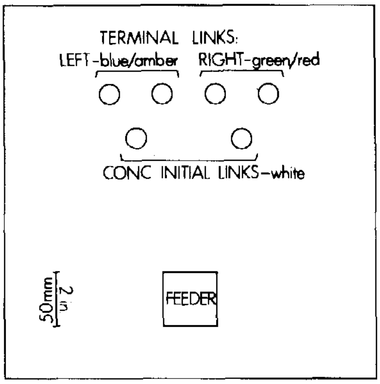


Fig. 1. Diagram of the six-key experimental panel. Concurrent (CONC) initial links operated on the two lower keys, lit white; terminal links operated on either the two upper left keys, lit blue and/or amber, or on the two upper right keys, lit green and/or red. (From Catania, 1975, Fig. 1; Copyright 1975 by and reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.)

Preferences were studied within the context of concurrent-chain schedules (Autor, 1969; Herrnstein, 1964): according to concurrent initial-link schedules, responses produced separately operating terminal links. An example is shown in Fig. 2. During initial links, at the top, the two lower keys were lit white and the other keys were dark. According to independent variable-interval (VI) schedules, pecks on each key produced their respective terminal links, shown in brackets. The left bracketed conditions show a free-choice terminal link: with fixed-interval (FI) reinforcement available on either of two lit keys (blue and amber), a peck on either lit key produced food at the end of the FI. The right bracketed conditions show a forced-choice terminal link: with FI reinforcement available on only a single lit key (green), only a peck on that key could produce food at the end of the FI. Each terminal link ended with the delivery of a single reinforcer, after which the initial links were reinstated. With this arrangement, relative rates of pecking on the two initial-link keys correspond to preferences for the respective free-choice and forced-choice terminal links.

Concurrent-chain schedules separate the preference for different conditions (in initial links) from the contingencies that maintain responding in those conditions (in terminal links). This is not the case with concurrent schedules that simply arrange different contingencies for two responses; if one response is maintained at a higher rate than the other, this does not necessarily imply that the first response is preferred to the second. For example, if a fixed-ratio (FR) schedule maintained higher response rates than a concurrent differential-reinforcement-of-low-rate (DRL) schedule, it would be inappropriate to conclude that FR responding is preferred to DRL responding. Pairs of concurrent schedules can be arranged concurrently (e.g., Menlove, Moffitt, & Shimp, 1973), but the contingencies controlling the separate responses and the changeovers among them may override the relative magnitudes of the reinforcers correlated with each schedule (cf. Leigland, 1979). The advantage of concurrent-chain schedules is that they do not confound preferences with the contingencies that operate on the distribution of responses among alternatives.

In the present procedures, concurrent VI 30-sec VI 30-sec schedules operated in initial links. Each schedule was made up of 20 intervals, constructed according to the specifications of Catania and Reynolds (1968, Appendix II). The schedules were identical, but operated independently and in opposite directions. The intervals of each schedule were timed from the end of reinforcement in the preceding terminal link of that schedule; neither schedule operated during terminal links or during reinforcement. Daily sessions ended after 15 min of concurrent initial links. Through the first 609 sessions, any initial link peck, including the first after a changeover from the other key, was eligible to produce a scheduled terminal link; thereafter, the first peck after a changeover from the other

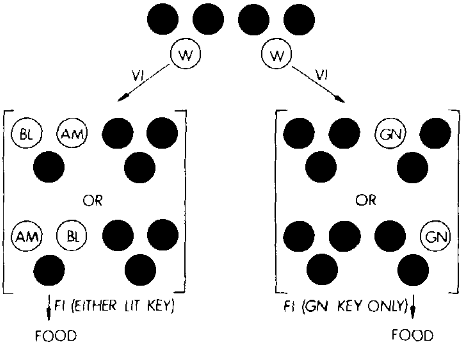


Fig. 2. Example of the concurrent-chain procedures. During initial links (top), the upper four keys were dark and the two lower keys were lit white (W). According to independent concurrent variable-interval (VI) schedules, pecks on these keys produced terminal links. In the left terminal link (left brackets), the two upper left keys were lit either blue (BL) and amber (AM) or amber and blue; the positions of the two colors changed irregularly over successive terminal links. In the right terminal link (right brackets), one of the two upper right keys was lit green (GN); its position changed irregularly over successive terminal links. In both terminal links, a peck on a lit key produced food at the end of a fixed interval (FI) and the initial links were then reinstated. In this example, left terminal links provided blue-amber free choices, and right terminal links provided a forced choice of green. (From Catania, 1975, Fig. 2; copyright 1975 by and reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.)

initial-link key was ineligible to produce a terminal link (changeover ratio or COR 2).

In terminal links, FI 20-sec schedules were arranged, operating for pecks on the single lit key in forced-choice terminal links and for pecks on either lit key in free-choice terminal links. With two lit keys, changeovers between the keys had no effect; the first peck after 20 sec was reinforced without regard to the prior sequence of pecks on the two keys during the 20-sec interval. All terminal links ended after a single FI reinforcement; thus, a peck on only one of the two keys could be reinforced within any given free-choice terminal link. The stimuli and schedules assigned to the two keys alternated irregularly over successive terminal links (e.g., as illustrated in Fig. 2 by the two arrangements of key colors in the free-choice terminal link, and by the two positions of green in the forced-choice terminal link).

The various free-choice and forced-choice conditions arranged in left and right terminal links are illustrated in Fig. 3. The top half of the figure shows cases in which forced choices are limited to a particular terminal-link color; the bottom half shows those in which forced choices are limited to a particular terminal-link key position. The left half of the figure shows cases in which free choice was arranged in left terminal links; the right half shows those in which free choice was arranged in right

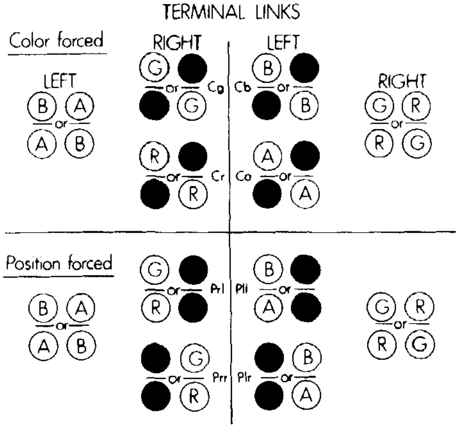


Fig. 3. Summary of free-choice and forced-choice terminal-link conditions. The upper left shows free choice of blue (B) and amber (A) in left terminal links and forced choice of green (top) or red (bottom) in right terminal links; the upper right shows free choice of green (G) or red (R) in right terminal links and forced choice of blue (top) or amber (bottom) in left terminal links. The lower left and right show corresponding arrangements with position rather than color forced in forced-choice terminal links. The key positions shown above and below the horizontal line within each terminal-link condition alternated irregularly. (Adapted from Catania, 1975, Fig. 3.)

terminal links. One sequence of procedures examined preference for free choice over forced choice with both color forced (sessions 164 to 198 and 238 to 244) and position forced (sessions 199 to 237); a second sequence examined preference only with color forced (sessions 641 to 678). Conditions were ordered so that those with free choice in the left terminal link alternated with those with free choice in the right terminal link. Conditions were typically maintained for seven to ten consecutive daily sessions (cf. Catania, 1975), and data were arithmetic means over the last three sessions of each condition.

The several forced-choice conditions followed from the assumption that any observed preference for free-choice terminal links would be difficult to assess, and perhaps uninteresting, if it were small relative to color or position preferences (for example, a preference might depend on how a particular terminal-link key favored observation of and movement toward the feeder). Free-choice preference was therefore examined with forced choice arranged for each possible color and position in the other terminal link.

B. PREFERENCES WITHIN CONCURRENT-CHAIN SCHEDULES

Preferences for free-choice over forced-choice terminal links are summarized for each pigeon in Fig. 4. Successive conditions are shown from top to bottom of the figure. In baseline (BL) sessions, equivalent forced-choice conditions were arranged in both terminal links. Each data point shows relative responding in the left initial link (left initial-link responses divided by left plus right initial-link responses). Apex-left triangles represent conditions with free choice in the left terminal link; apex-right triangles represent those with free choice in the right terminal link. The x-axis is scaled so that displacements to the left or right correspond respectively to shifts of preference to left or right terminal links. Thus, shifts in relative initial-link response rate that follow the directions in which the triangles point correspond to changes in preference that follow free-choice terminal links.

For Pigeon 53 (left), preference shifted to the left from its baseline value when free choice was introduced in the left terminal link (C_g). When free choice was then moved to the right terminal link (C_a), preference shifted to the right. Each subsequent change in the location of the free-choice terminal link was accompanied by a corresponding shift of preference. Similar shifts were obtained in the second sequence (bottom connected points), after more than 400 sessions of other procedures had intervened. The data for Pigeons 211 (middle) and 280 (right), though

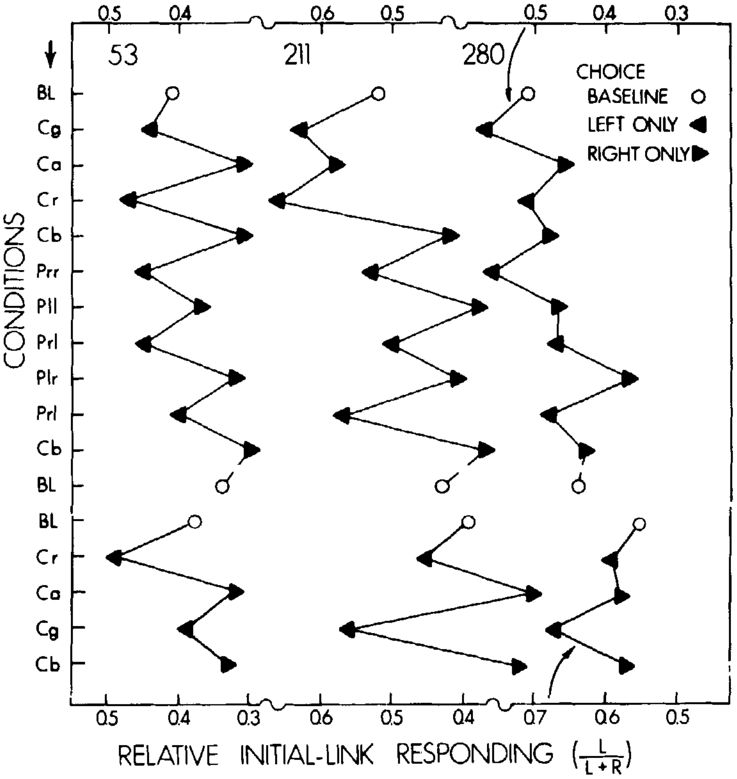


Fig. 4. Relative initial-link responding over successive free-choice and forced-choice conditions for three pigeons. Conditions, shown from top to bottom in the order in which they were studied, included baseline (BL), free choice versus forced choice of color (C), and free choice versus forced choice of position (P); the notation along the *y*-axis corresponds to that of Fig. 3. Free-choice versus forced-choice data with free choice in the left terminal link are shown by apex-left triangles; those with free choice in the right terminal link are shown by apex-right triangles. The *x*-axis has been scaled so that displacements to the left indicate increased left terminal-link preference and displacements to the right indicate increased right terminal-link preference; thus, a displacement corresponding to the direction in which a triangle points implies a shift in preference toward a free-choice terminal link. Different scales are provided for each pigeon and for the two successive sets of data for Pigeon 280; the upper and lower sets of data were separated by about 400 sessions. Each point represents the arithmetic mean over the last three sessions of a condition. (From Catania, 1975, Fig. 4; copyright 1975 by and reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.)

different in absolute magnitude, were comparable in direction. The only exception was one shift of the free-choice terminal link from right to left (P_{LL} to P_{RL}) for Pigeon 280; in this instance, the relative rates in the two conditions were equal to two decimal places, and did not constitute a reversal. Thus, in 44 of the 45 schedule changes in Fig. 4, relative rates shifted in accordance with preference for free-choice terminal links. The shifts in preference were typically smaller for Pigeon 280 than for the other two pigeons; the individual differences were consistent through both sequences of procedures.

The data presentation separates shifts in free-choice preference from shifts in baseline preference, when equivalent conditions were arranged in the two terminal links. For Pigeon 53, the baseline relative rate remained in the region of .4 throughout both sequences. For Pigeons 211 and 280, the baseline relative rate began in the region of .5 and moved to about .4 by the end of the first sequence. For Pigeon 211, the baseline shift occurred abruptly (conditions C_r to C_b); for Pigeon 280, the shift occurred gradually. In the second sequence, the baseline relative rate remained at about .4 for Pigeon 211, but had shifted almost to .6 for Pigeon 280, as shown by the different top and bottom scales.

Cumulative records over a full session of both initial-link and terminal-link responding are shown for each pigeon in Fig. 5. The lower-right inset shows initial-link relative-rate records from the same session (the latter records were obtained with a standard Gerbrands cumulative recorder, using vertical response steps for right initial-link responses and horizontal .2-sec operations of the paper-drive motor for left initial-link responses; cf. Kulli & Bogrow, 1971). In these records, positive curvature indictes a decreasing left terminal-link preference over successive initial-link responses, and negative curvature indicates an increasing left terminal-link preference. The inset therefore demonstrates that, perhaps because of local asymmetries in relative reinforcement for the two initial links, substantial changes in preference occurred within sessions; these changes did not necessarily correspond in direction to session-to-session changes in preference.

The session illustrated was the first session with free choice in the right terminal link after sessions of free choice in the left terminal link (C_r to C_a, bottom of Fig. 4). For the three pigeons, preferences were respectively .49, .45, and .59 over the last three sessions of the previous condition (free choice on left), .44, .36, and .58 in the session of Fig. 5 (free choice on right), and .32, .30, and .58 over the last three sessions of the latter condition. In general, shifts in preference were observed within the first session after a change in terminal links, and the major part of the shift in preference produced by each change was typically complete within two or three sessions. For example, the patterns of results throughout the present experiments would not have been different if means over the last three of the first five sessions of each condition had been substituted for those over the last three of all sessions for that condition.

The experiments of Fig. 4 could have been conducted in a standard

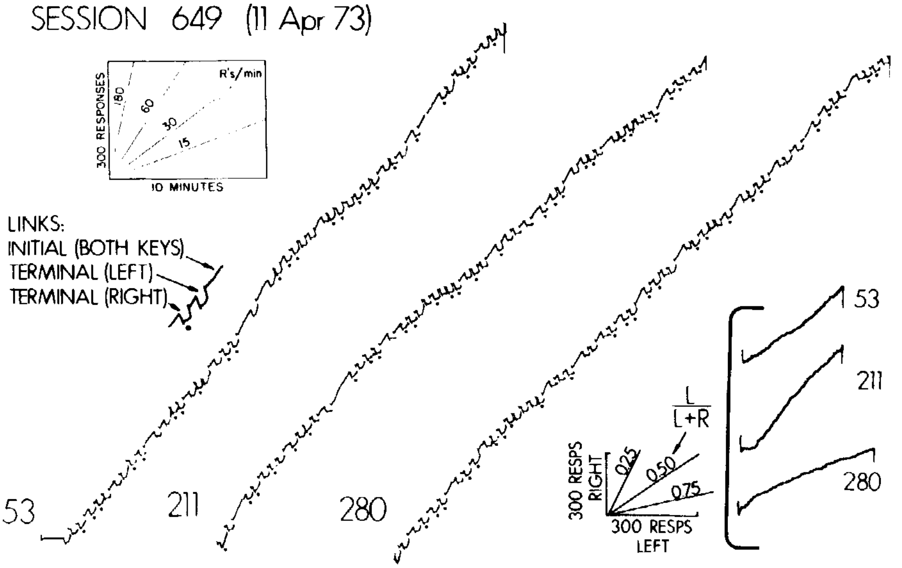


Fig. 5. Cumulative records of full sessions of initial-link and terminal-link performance for each pigeon in the first session of right terminal-link free choice after sessions of left terminal-link free choice (third from last condition in Fig. 4). The recording pen was displaced downward during both terminal links; right terminal links are distinguished by dots under the record. The inset at lower right shows relative-rate records from the same session. Each left initial-link response stepped the recording pen vertically, and each right initial-link response stepped it horizontally (.2-sec operations of the paper-drive motor; cf. Kulli & Bogrow, 1971). Left and right step-sizes were unequal, and slopes for some representative relative rates are shown in the accompanying scale.

two-key pigeon chamber; in these procedures, no more than two of the six keys were ever lit at the same time. The spatial separation of initial links, left terminal links, and right terminal links may have contributed to the rapid adjustment of initial-link responding to changes in terminal links and to the stimulus control exerted by the terminal-link colors. The several invariant features of the terminal-link colors and key positions were chosen on the assumption that they would enhance the stability and sensitivity of the concurrent-chain performance. Although no experimental comparison of the six-key procedure with standard two-key concurrent-chain schedules was conducted, the present arrangement allowed briefer exposure to each condition than has typically been used with these schedules (e.g., Duncan & Fantino, 1972; Fantino, 1968; Herrnstein, 1964; Moore & Fantino, 1975; Navarick & Fantino, 1972).

In the standard records of Fig. 5, left terminal links (forced choice) are shown by unmarked pen displacements, and right terminal links (free choice) by pen displacements marked with dots. The temporal pattern of responding varied over successive instances of both types of terminal links. Forced-choice terminal links include more instances of abrupt transitions from a pause to a relatively high FI rate than free-choice terminal links in the sessions shown, but the range of pauses and scalloping overlapped considerably across the two types of terminal link; over sessions, no consistent relation between temporal patterning and preference was evident from visual inspection of the records.

III. Analyzing the Free-Choice Preference

> Clearly there can be no simple or direct relationship between the range of available alternatives and the extent of freedom. However numerous the alternatives between which a man may choose, he will not admit himself to be free if the one alternative that he would most prefer is the one which is excluded.
>
> P. H. PARTRIDGE

> There's small choice in rotten apples.
> WILLIAM SHAKESPEARE

In the preceding experiment (Fig. 4), each pigeon preferred terminal links that provided two alternative routes to a food reinforcer (free choice) to those that provided only a single route to the same reinforcer (forced choice). But what was the basis of the preference? Was it determined directly by the number of alternatives, or was it determined indirectly, in that the alternatives produced differences in terminal-link responding that affected preference in turn? Was the preference merely a function of the number of stimuli, and not of the correlation of those stimuli with alternative routes to the reinforcer? Even if the availability of a choice was critical, what constituted a choice? If dark keys ought to be counted as choices, then every procedure in the present apparatus involved six alternatives; for that matter, it could be argued that an indefinite number of alternatives was always available in the indeterminate number of differentiable locations at which the pigeon might have pecked. This section considers these issues by examining the relations among preferences, properties of terminal-link performances, and stimulus variables within the concurrent-chain schedules. Both correlational and experimental analyses of concurrent-chain performance may be relevant to our understanding of the free-choice preference.

A. FREE-CHOICE PREFERENCE AND TERMINAL-LINK PERFORMANCE

The available literature does not suggest that the rate or pattern of terminal-link responding is likely to affect initial-link preferences within the present concurrent-chain procedures (e.g., Fantino, 1968; Killeen, 1968; Moore & Fantino, 1975; Neuringer, 1969). Nevertheless, the possibility must be considered that preferences depended on different properties of free-choice and forced-choice terminal-link responding. One variable of potential interest is relative terminal-link responses per reinforcement or, equivalently (because equal FI schedules were arranged in each terminal link), relative terminal-link response rates. To the extent that factors that generate higher response rates in one terminal link than the other also favor preference for the former over the latter terminal link, initial-link and terminal-link response rates will be positively correlated; on the other hand, to the extent that lower-rate terminal-link responding is preferred because it involves fewer responses per reinforcement, initial-link and terminal-link response rates will be negatively correlated.

Figure 6 shows scatterplots of relative initial-link and terminal-link response rates for each pigeon in the preceding experiment and in several related concurrent-chain procedures (cf. Catania, 1975, Figs. 4, 6, 8, and 9). Each experimental condition is represented by the mean relative response rates over the last three sessions. No correlation is obvious in these data; if a positive or negative correlation exists between relative initial-link and terminal-link response rates, it is too weak a one to be invoked in an account of free-choice preference.

One way in which terminal-link performance might have influenced preference was through its effect on time to reinforcement, because this time was equal to the 20-sec fixed interval plus the time from the end of this interval to the reinforced peck. Terminal-link response rates, however, were sufficiently high that the latter times were short, and relative times to reinforcement in free-choice and forced-choice terminal links

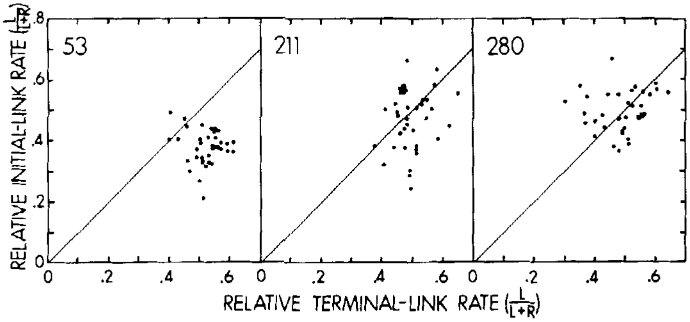


Fig. 6. Scatter-plot of relative initial-link rate as a function of relative terminal-link rate for each pigeon. The diagonal shows the locus of equal relative rates. (Adapted from Catania, 1975, Figs. 4, 6, 8, and 9.)

typically differed by less than 1%; thus, the effect of this variable on initial-link preference was presumably negligible.

Another aspect of terminal-link performance was the distribution of responses to the two alternatives during free choice. In no free-choice terminal link was responding restricted exclusively to one key color or to one key position, but the proportions of responding maintained by each terminal-link color and position did vary over a substantial range. Figure 7 shows relative initial-link rates for each pigeon as a function of the difference in relative response rates on the two free-choice terminal-link keys; data are shown in absolute value for both the two terminal-link colors (top) and the two terminal-link key positions (bottom). Consider, for example, the left terminal link of the procedure illustrated in Fig. 2. If blue and amber maintain equal terminal-link responding, this statistic is .0; if blue maintains nine-tenths of the responding and amber maintains the remaining tenth, this statistic is .8. No obvious relation between this statistic and relative initial-link rates is evident in Fig. 7, although the filled and unfilled data points might be expected to converge as the statistic approaches 1.0: the case in which all responding is restricted to one alternative could be interpreted as an instance of forced choice rather than free choice.

B. NUMBER OF LIT KEYS AND PREFERENCE FOR INFORMATIVE STIMULI

In each free-choice terminal link, two keys were lit; in each forced-choice terminal link, only a single key was lit. Even though both types of terminal links made reinforcers available equally often and according to the same schedule, preference might have depended on the presentation of two lit keys rather than a single lit key. Visual stimuli can serve as reinforcers, and some visual stimuli have a greater reinforcing effect than others (e.g., Berlyne, 1966; Munsinger, Kessen, & Kessen, 1964). For example, in concurrent schedules of stimulus presentation with humans, higher response rates were maintained by the more complex member of a pair of visual stimuli (Berlyne, 1972).

This problem was addressed by arranging concurrent-chain schedules in which some terminal links included lit keys not correlated with reinforcement (Catania, 1975, Fig. 5). The procedures are illustrated in Table I. For example, in procedure A (cf. Fig. 2), blue and amber in the left terminal link and green in the right terminal link are each correlated with reinforcement. The number of lit keys in these terminal links can be equated, as in procedure B, by presenting both red and green keys in the right terminal link but allowing pecks on only the green key to produce

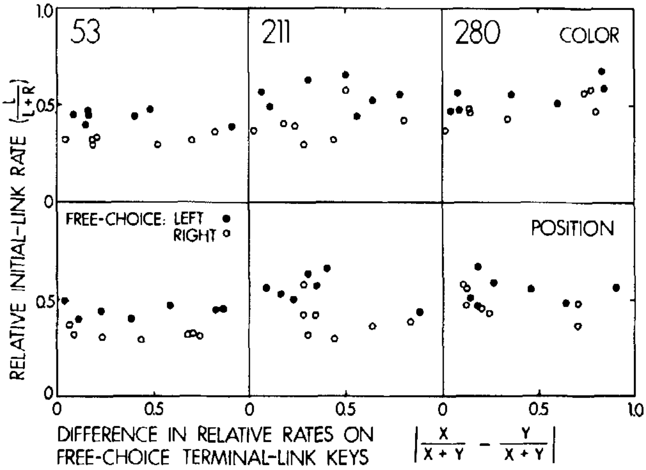


Fig. 7. Relative initial-link rate as a function of the absolute difference in relative rates within free-choice terminal links for each pigeon. Filled and unfilled circles respectively show data obtained with free choice in left and free choice in right terminal links. The upper graphs are based on the difference between terminal-link rates for keys of different colors; the lower graphs are based on the difference for keys in different positions.

the reinforcer. Procedure C allows pecks on blue but not amber to produce the reinforcer in left terminal links; in this case, the terminal links differ in number of lit keys, but each is a forced-choice condition because each allows only a single route to the reinforcer. These procedures, however, produced equivocal results: preference for two lit keys over one lit key was evident for only some pigeons in some procedures (Catania, 1975, Fig. 6). One source of the variable results may have been that the lit keys correlated with nonreinforcement were still maintaining responding during the sessions in which preferences were determined.

It might be argued that two lit keys provide a choice, albeit a highly determined one (cf. Hobson's choice), even if one of the keys is correlated with extinction (EXT). But if a lit key correlated with EXT constitutes an alternative, then an unlit key correlated with EXT should do so also, and procedures substituting lit EXT keys for unlit EXT keys should be indistinguishable. Such an interpretation reduces all terminal links in the present apparatus to a choice among six alternatives, and suggests that unlit keys should be ignored for the purposes of the present analysis. An additional justification is that the discrimination between lit and unlit keys had been established by an extensive experimental history, and pecks on unlit keys were rare during experimental sessions (probably because at least one lit key was always available).

Given the distinction between lit and unlit keys, another possible function of lit keys must be considered. If reinforcement and extinction are arranged for each of two keys, then the stimuli correlated with these schedules are informative (a stimulus that specifies one of the two keys provides exactly one bit of information). A concurrent-chain procedure designed to vary the number of lit keys in terminal links while holding the number of alternatives constant will therefore necessarily affect the informative functions of the stimuli (cf. Table I, B). For this reason, studies of preferences for informative over uninformative stimuli were included among the present experiments.

Research on the reinforcing effect of informative stimuli constitutes an extensive and controversial literature (e.g., Egger & Miller, 1962; Bower, McLean, & Meacham, 1966; Gollub, 1970). One difficulty in such studies is that of separating the presentation of different stimuli from the different distributions in time of the reinforcers delivered in the presence of each. With concurrent chains, an advantage of scheduling FI reinforcement for pecks on one or both of two terminal-link keys is that the stimuli may be varied without affecting the time at which the reinforcer becomes available in the presence of each stimulus: informative stimuli can be added to terminal links without altering their relative frequencies of reinforcement.

Some concurrent-chain terminal links involving informative stimuli are illustrated in Fig. 8. Each terminal link includes two lit keys; FI reinforcement is scheduled for one key but not the other, and the position of this key alternates irregularly over successive terminal links. In one case (upper left), no informative stimuli are provided in either terminal link; in both left and right terminal links, the two keys are each lit the same color. Informative stimuli can be introduced in the left terminal link by lighting the FI key blue and the EXT key amber (upper right); similarly, informative stimuli can be introduced in the right terminal link by lighting the FI

TABLE I

SOME TERMINAL-LINK ARRANGEMENTS FOR THE STUDY OF PREFERENCE FOR TWO LIT KEYS OVER ONE LIT KEY

Procedure	Left terminal-link keys[a]	Right terminal-link keys[a]
A. Free choice versus forced choice, two lit keys versus one lit key	Blue (FI), amber (FI)	Green (FI), dark (EXT)
B. Free choice versus forced choice, two lit keys versus two lit keys	Blue (FI), amber (FI)	Green (FI), red (EXT)
C. Forced choice versus forced choice, two lit keys versus one lit key	Blue (FI), amber (EXT)	Green (FI), dark (EXT)

[a] FI, fixed interval; EXT, extinction.

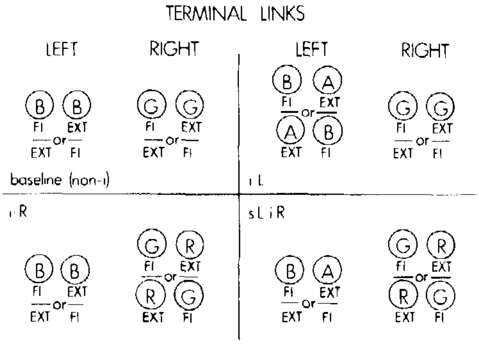


Fig. 8. Some procedures involving informative stimuli. In each terminal link, pecks on one key produce food at the end of a fixed interval (FI), and pecks on the other key have no effect (EXT); these contingencies alternate irregularly between the two keys. The upper left illustrates a baseline condition in which no informative stimuli are provided in either terminal link: both left terminal-link keys are blue (B), and both right terminal-link keys are green (G). In the upper right, informative stimuli are provided in the left terminal link: the FI key is blue and the EXT key is amber (A). In the lower left, informative stimuli are provided in the right terminal link: the FI key is green and the EXT key is red (R). The lower right shows varied stimuli in both terminal links, but informative stimuli in only the right terminal link; the blue and amber left terminal-link stimuli are not correlated with the FI and EXT contingencies for the two keys. (Adapted from Catania, 1975, Fig. 7.)

key green and the EXT key red (lower left). In both of the conditions, two same-color stimuli are presented in one terminal link and two different-color stimuli are presented in the other. Thus, the effects of these two conditions were compared with those of concurrent-chain procedures in which FI reinforcement was scheduled for both terminal-link keys and in which the terminal links differed only in the variety of the key colors, i.e., two same-color keys versus two different-color keys. (A procedure in which two different-color keys are provided in both terminal links but in which the FI and EXT keys are correlated with colors in only one terminal link is illustrated in the lower right panel of Fig. 8; this procedure was not used in the present study.)

Figure 9 summarizes the preferences obtained with both the informative-stimulus (INF) and the variety (V) procedures. Unfilled apex-left triangles show data obtained with informative stimuli in the left terminal link; unfilled apex-right triangles show data obtained with informative stimuli in the right terminal link. Baseline (BL) data obtained with informative stimuli in neither terminal link are shown by filled circles. Similarly, data for variety (two different-color keys) in left and in right terminal links are shown respectively by apex-left and apex-right filled triangles; baseline data for this procedure are shown by unfilled circles. As in Fig. 4, successive conditions are shown from top to bottom, and displacements to the left or right correspond in direction to changes in preference for left or right terminal links.

For each pigeon, relative initial-link rates shifted in a direction consistent with preferences for informative stimuli with each change of conditions. The exception was the last change to baseline from informative stimuli in the left terminal link for Pigeon 280 (as for the corresponding case in Fig. 4, the relative rates, equal to two decimal places, did not constitute a reversal). Thus, 17 of the 18 schedule changes demonstrated a preference for informative over uninformative stimuli. The approximate baseline levels and the ordering of the three pigeons in the magnitude of preference were comparable to those of Fig. 4.

With respect to stimulus variety, preferences were unsystematic for Pigeon 53 and were small relative to those for informative stimuli for Pigeons 211 and 280. One statistic for measuring magnitude of preference for a given feature of terminal links is one-half the difference between relative rates with that feature in the left terminal link and those with that feature in the right terminal link; this statistic expresses preference relative to the baseline (equivalent left and right terminal links). Averaged across the three pigeons, the preference for free choice over forced choice (from Fig. 4) was .06, that for informative stimuli was .08, and that for

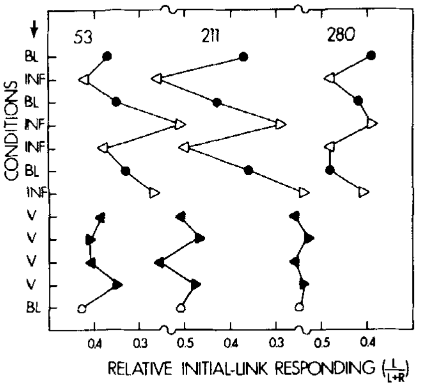


Fig. 9. Relative initial-link rates over successive conditions in procedures arranging stimulus information (INF) and stimulus variety (V) in terminal links. Baseline (BL) data are shown by filled and unfilled circles. Data for informative stimuli in left terminal links are shown by apex-left unfilled triangles; those for informative stimuli in right terminal links are shown by apex-right triangles. The procedures correspond respectively to the upper right and lower left conditions in Fig. 8. Data for varied stimuli in left or in right terminal links are shown respectively by apex-left and apex-right filled triangles. The left and right terminal-link stimuli in the successive variety conditions were respectively blue-amber versus green-green, blue-blue versus green-red, blue-amber versus red-red, and amber-amber versus green-red. (From Catania, 1975, Fig. 8; copyright 1975 by and reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.)

stimulus variety was less than .02 (although the data were variable, the equivalent statistic for two lit keys versus one lit key was also about .02; cf. Table I, C, and Catania, 1975, Table 4). Given the comparable magnitudes of the free-choice and informative-stimulus preferences, experiments designed to study free-choice preference while controlling for number of lit keys must also be able to take the informative functions of the lit keys into account. For example, with free choice of blue and amber keys in the left terminal link and forced choice of a green key in the right terminal link (see Fig. 2 and Procedure A of Table I), adding a red EXT key to the right terminal link would make these terminal links equal with respect to a number of lit keys (see Procedure C of Table I), but would also pit free choice in the left terminal link against informative stimuli in the right terminal link.

C. FREE-CHOICE PREFERENCE AND TERMINAL-LINK STIMULUS VARIABLES

The preceding research demonstrated the pigeon's preference for free choice over forced choice, but controlled for stimulus number, stimulus information and stimulus variety only indirectly, across a sequence of procedures. The following experiment (Catania & Sagvolden, 1980) more directly separated free-choice preference from the effects of these other variables. Concurrent chains were again scheduled in the apparatus of Fig. 1. Both free-choice and forced-choice conditions were arranged in terminal links on the four top keys; an illustration of the procedure is provided in Fig. 10. In the forced-choice condition (left terminal link in Fig. 10), pecks on a single green key could produce FI food reinforcement while pecks on the three remaining red keys had no scheduled consequences. In the free-choice condition (right terminal link in Fig. 10), pecks on any one of the three green keys could produce FI reinforcement while pecks on the remaining red key had no scheduled consequences. Thus, the two conditions were equated with respect to stimulus number (four lit keys versus four lit keys), stimulus variety (one odd color among four versus one odd color among four), and traditional measures of information (two bits of information to specify one of four keys versus two bits of information to specify three of four keys).

The pecking of four male White Carneaux pigeons, about 1 year old at the start of the research and maintained at about 80% of free-feeding body weights, had been established in a two-key autoshaping procedure in which one key was red and the other was green (Fisher & Catania, 1977); in that procedure, each pigeon had shown a red-key preference. Daily sessions of concurrent chains were ordinarily arranged for 25 min of

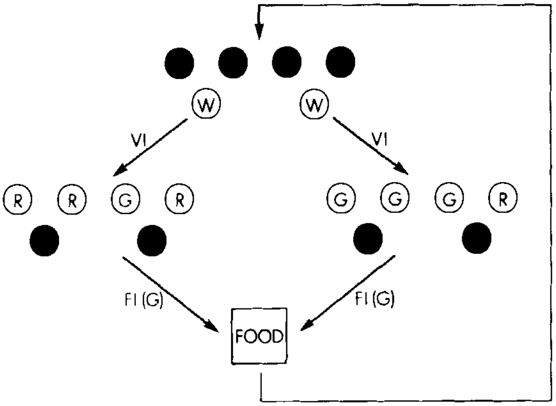


Fig. 10. Example of the concurrent-chain procedure. During initial links, the top four keys were dark and the two bottom keys were lit white (W). According to independent variable-interval (VI) schedules, pecks on the left and right white keys produced their respective and mutually exclusive terminal links. During terminal links, the top keys were lit green (G) and red (R) and the bottom two keys were darkened. In the example, a forced-choice terminal link is illustrated on the left, and a free-choice terminal link on the right. In forced-choice terminal links, one key was lit green and the three remaining keys were lit red. In free-choice terminal links, three keys were lit green and the one remaining key was lit red. In both terminal links, a peck on any green key produced food at the end of a fixed interval (FI); a peck on a red key had no effect. After each food delivery, initial-link conditions were reinstated. (From Catania & Sagvolden, 1980, Fig. 1; copyright 1980 by and reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.)

initial links, but were occasionally increased or decreased in 5-min steps to maintain 80% weights while holding post session feeding to a minimum; session durations therefore ranged from 20 to 35 min. Each key operated at a force of about .14 N, and pecks on lit keys produced feedback clicks from a relay mounted behind the panel.

In these procedures, independent concurrent VI 30-sec schedules were arranged for pecks on the two white bottom keys during initial links; the top four keys were dark. The VI schedules each consisted of 15 intervals constructed according to the specifications of Catania and Reynolds (1968, Appendix II). As in the preceding experiments, the timing of intervals stopped for a given initial-link key when a terminal link had been scheduled for the next peck on that key, and the timing of intervals stopped for both keys during terminal links and food deliveries. The first peck after a changeover from one initial-link key to the other was not eligible to produce a terminal link (changeover-ratio 2). In terminal links, the top four keys were lit and the bottom two keys were dark. In both free-choice and forced-choice terminal links, pecks on any green key were reinforced according to an FI 30-sec schedule. After reinforcement, a 3-sec food delivery during which the feeder was lit and all keys were dark, the initial links were reinstated. Throughout all conditions, pecks on dark keys had no scheduled consequences.

In most sessions, the position of the odd-colored key varied unsystematically over successive terminal links. With top keys numbered from left to right, Fig. 10 shows green located on Key 3 in the forced-choice terminal link (left), and red located on Key 4 in the free-choice terminal link (right). The effects on preference of terminal-link response rates and stimulus control were examined in some sessions by fixing the location of the odd-colored terminal-link key. In both free-choice and forced-choice terminal links, Key 1 maintained the highest green response rates and Key 3 the lowest green response rates for most pigeons (Catania & Sagvolden, 1980, Table I). In one procedure, Key 1 was always green and the remaining keys were red in forced-choice terminal links, whereas Key 1 was always red and the remaining keys were green in free-choice terminal links. In another procedure, Key 3 was always green and the remaining keys were red in forced-choice terminal links, whereas Key 3 was always red and the remaining keys were green in free-choice terminal links.

The sequence of conditions and the daily sessions of each condition are combined with the presentation of data in Fig. 11. As in Figs. 4 and 9, successive conditions are presented from top to bottom along the y-axis, the x-axis is scaled so that a shift to the left corresponds to an increase in preference for the left terminal link and one to the right corresponds to an increase in preference for the right terminal link, and the data points are apex-left triangles for free choice in left terminal links and apex-right triangles for free choice in right terminal links. For all four pigeons, preferences shifted from right to left when free choice was switched from right to left terminal links (after sessions 8 and 108) and from left to right when free choice was switched from left to right terminal links (after session 51). Compared to this main effect, relatively small and inconsistent effects were produced by manipulating the positions of the green and red terminal link keys. The two largest of these (third condition for Pigeon 18 and eighth condition for Pigeon 23) occurred with responding forced to the high-rate terminal-link green key in forced-choice terminal links, but were in opposite directions; the magnitude of free-choice preference decreased for Pigeon 18 and increased for Pigeon 23. No consistent effects of key position are evident in the mean data shown on the right in Fig. 11.

When positions were fixed so that Key 1 was always red in free-choice terminal links and green in forced-choice terminal links, a preferred key position was restricted to the forced-choice terminal link. When positions were fixed so that Key 3 was always red in free-choice terminal links and green in forced-choice terminal links, the more preferred key positions

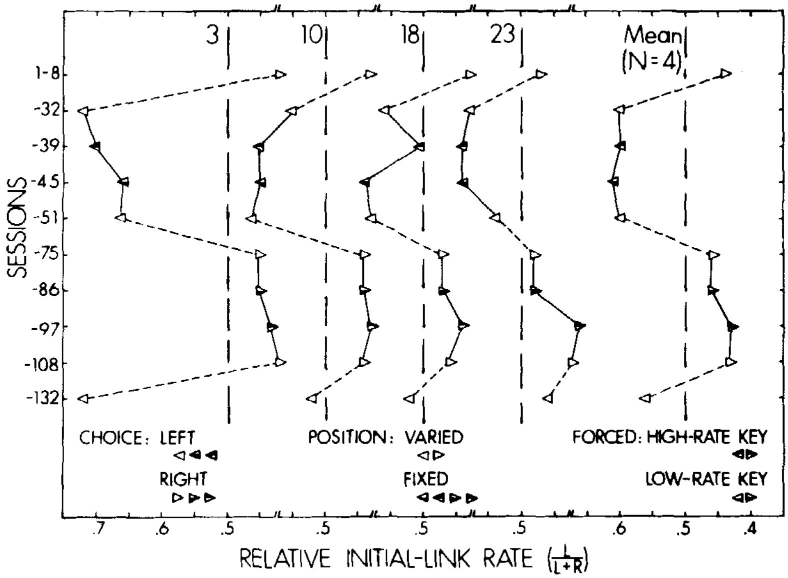


Fig. 11. Relative initial-link response rates over the last five sessions of each condition for each of four pigeons. Data with free choice in left terminal links are shown by apex-left triangles and those with free choice in right terminal links are shown by apex-right triangles. Free-choice terminal links included three green keys and one red key; forced-choice links included one green key and three red keys (cf. Fig. 10). The positions of red and green keys in successive terminal links were either varied or fixed. When fixed, forced-choice responding (single green key) was limited either to a formerly high-rate key or to a formerly low-rate key. (From Catania & Sagvolden, 1980, Fig. 3; copyright 1980 by and reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.)

were restricted to the free-choice terminal link, but a key position that might have been correlated with a lower ratio of responses to reinforcers by virtue of its lower response rate was restricted to the forced-choice terminal link. In neither case was there a systematic effect on the magnitude of free-choice preference. As in the preceding studies, relative terminal-link response rates and the distribution of responses among the terminal-link keys were also not correlated with changes in preference (Catania & Sagvolden, 1980, Fig. 4 and Table I).

Another property of terminal-link responding that might have been relevant to preference was stimulus control by the green and red keys. Only one red key was present during free-choice terminal links, whereas three red keys were present during forced-choice links. Even in the absence of stimulus control by red and green, the probability of pecking a red (EXT) key was only .25 in free-choice terminal links, whereas it was .75 in forced-choice terminal links. Terminal-link red-key responding is summarized in Table II, which shows that, although the rate of red-key pecking was low throughout, under most conditions less red-key pecking occurred in free-choice terminal links (one red key) than in forced-choice terminal links (three red keys). In seven instances, however, this relation was reversed; of the reversals, five occurred when the red key in free-choice terminal links was fixed at the high-rate (Key 1) position. The reversals were not systematically related to changes of preference.

This study replicated the pigeon's preference for free choice within concurrent-chain schedules even though the terminal-link stimuli did not differ in number, variety, or information (in bits). Furthermore, the preference was not correlated with terminal-link position preferences, relative terminal-link response rates, or the degree of discriminative control exerted by the terminal-link stimuli. But in any study, the preference for free choice will inevitably be confounded with other variables. Alternatives cannot be presented without accompanying stimuli, and these stimuli set the occasion for responding. Thus, the availability of alternatives might be spoken of in terms of presenting two or more conditioned reinforcers, or of setting the occasion for two or more discriminated operants, or of providing the opportunity for two or more highly probable

TABLE II

RATES OF PECKING (RESP/MIN) ON RED (EXT) KEYS IN FREE-CHOICE AND FORCED-CHOICE TERMINAL LINKS[a]

	Pigeon							
	3		10		18		23	
Terminal-link choices	Free	Forced	Free	Forced	Free	Forced	Free	Forced
Free-choice-left								
Varied position, 3G1R-1G3R	.01	.06	.09	1.13	.02	.12	.01	.03
Fixed position, RGGG-GRRR	.02	.02	.46	.24	.11	.04	.08	.00
Fixed position, GGRG-RRGR	.00	.17	.00	2.25	.00	.00	.00	.07
Varied position, 3G1R-1G3R	.01	.10	.34	2.36	.00	.01	.00	.04
Free-choice-right								
Varied position, 1G3R-3G1R	.03	.12	.26	1.09	.00	.00	.30	.03
Fixed position, RRGR-GGRG	.03	.12	.03	2.93	.07	.00	.10	.45
Fixed position, GRRR-RGGG	.00	.10	.14	.10	.00	.01	.22	.03
Varied position, 1G3R-3G1R	.00	.14	.16	.46	.00	.00	.02	.14

[a] Data are arithmetic means over the last five sessions of each condition for each pigeon. With varied position, free choice (three green, one red) is represented as 3G1R and forced choice (one green, three red) as 1G3R. With fixed position, the orders of green (G) and red (R) correspond to the four top key positions.

responses. These usages respectively emphasize stimuli, stimulus–response relations, and responses, but the priorities among them are a matter of technical vocabulary rather than of the empirical status of free-choice preference.

For example, consider the language of response variety: a single alternative implies the availability of only one highly probable response, whereas two or more alternatives imply the availability of two or more highly probable responses. Preference may then be interpreted in terms of the summation of the latter response probabilities, even though the responses cannot be emitted simultaneously once they are available (cf. Premack, 1965, 1971). Yet how can the availability or unavailability of these responses have any behavioral effect unless the availability is correlated with some stimulus? The free-choice preference might then be restated in terms of a preference for an opportunity for more varied responding, without regard to the number of alternatives (cf. Boren, Moerschbaecher, & Whyte, 1978; Eckerman & Lanson, 1969; Herrnstein, 1961; Platt, 1973; Schoenfeld, Harris, & Farmer, 1966). In this view, the pigeon's free-choice preference is equivalent simply to preference for a larger area of pecking. Nevertheless, even the availability of different areas for pecking must be correlated with discriminative stimuli, and the empirical questions again involve the boundaries of those response classes called discriminated operants (cf. Catania, 1973). One empirical property of the free-choice preference that may bear on these several vocabularies and interpretations is the function relating magnitude of the free-choice preference to number of free-choice alternatives; this function is examined in the next section.

IV. Preference and Number of Free-Choice Alternatives

> Nor does it help to offer the principle of maximization of choice as the key to freedom. We must consider the possibility . . . that choice may become overchoice, and freedom unfreedom.
>
> ALVIN TOFFLER

During the 95 sessions that followed those in Fig. 12, key-pecking was reestablished on red terminal-link keys and Pigeon 10 was dropped from the experiment because of illness, low initial-link response rates, and highly variable performance. The concurrent-chain procedure was changed so that the upper four keys of the apparatus in Fig. 1 were lit only red in left terminal links and only green in right terminal links. Successive conditions then varied the number of lit keys, from one to four, in each

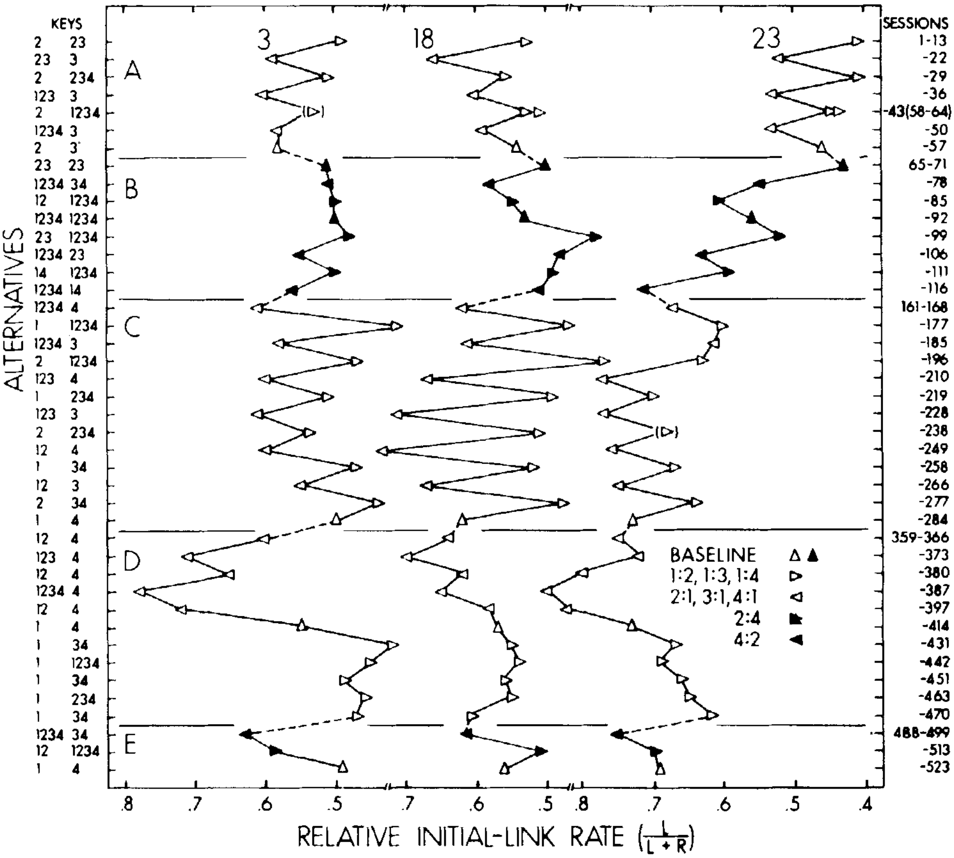


Fig. 12. Relative initial-link response rates with varying numbers of free-choice alternatives. Successive conditions are ordered along the y-axis, with the sessions at each condition shown on the right. Each condition is summarized on the left in terms of the keys respectively available in left and in right terminal links. In left terminal links, available keys were lit red; in right terminal links, available keys were lit green; in either terminal link, the remaining keys were dark, and a peck on any lit key produced food at the end of a fixed interval. The four terminal-link keys (top row, Fig. 1) are numbered from left to right. Thus, in the first condition, Key 2 was lit red in the left terminal link, and Keys 2 and 3 were lit green in the right terminal link. Baseline conditions (equivalent terminal links) are shown by apex-up triangles. Terminal-link conditions differing in the number of available keys are shown by apex-left and apex-right triangles. For conditions with free choice in one terminal link and forced choice in the other (unfilled triangles), the triangle points in the direction of the free-choice terminal link. For conditions with free choice in both terminal links (filled triangles), the triangle points in the direction of the terminal link with the larger number of alternatives. The x-axis is arranged as in Figs. 4, 7, and 11. The bracketed data point for Pigeon 3 in sequence A was obtained during a repetition of the condition after an apparatus problem for this pigeon; both sets of data are shown for Pigeons 18 and 23. At the bracketed data point for Pigeon 23 in sequence C, the last two sessions of the condition were excluded because of an apparatus problem. All data points are arithmetic means over the last three sessions of each condition.

terminal link; pecks on any lit key produced food reinforcers according to an FI 20-sec schedule, and pecks on dark keys had no scheduled consequences. As in preceding conditions, independent concurrent VI 30-sec schedules operated on the two bottom white keys during initial links, and the first peck after a changeover from one initial-link key to the other was not eligible to produce a terminal link (changeover ratio 2). During the first 116 sessions of the new procedure, reinforcement duration was 4 sec; thereafter it was reduced to 3.5 sec. Session durations, determined as for the procedures of Fig. 11, ranged from 20 to 30 min of initial links for Pigeons 3 and 18 and from 15 to 30 min of initial links for Pigeon 23.

Data from each of five sequences of conditions are presented in Fig. 12. The lit keys in the respective left and right terminal links are shown along the left *y*-axis; the sessions of each condition are shown along the right *y*-axis. As in Figs. 4, 9, and 11, the *x*-axis is arranged so that displacements to the left or right indicate corresponding shifts in preference. Conditions in which a single lit key was arranged in one or both terminal links are represented by unfilled data points; these include forced-choice baseline conditions and free-choice versus forced-choice conditions in which the number of free-choice alternatives was varied. Conditions in which two or more lit keys were arranged in both terminal links are represented by filled data points; these include baseline conditions with equal numbers of lit keys in each terminal link and conditions with four lit keys in one free-choice terminal link and two in the other. Apex-up triangles represent baseline conditions; apex-left and apex-right triangles point respectively in the direction of the terminal link with the greatest number of lit keys.

In Sequence A, the preference for free choice over forced choice was examined with two, three, or four lit keys in the free-choice terminal link. In left forced-choice terminal links Key 2 was lit, in right forced-choice terminal links Key 3 was lit, the location of the forced-choice terminal link alternated between left and right over successive conditions, and the fifth condition was repeated because of an apparatus failure. Sequence B examined preference for two versus four lit keys in terminal links. After 54 sessions of other procedures, Sequence C again examined the preference for free choice over forced choice with varying numbers of lit keys in the free-choice terminal link. Although the preceding research (Fig. 11) had suggested that key position in terminal links was not a critical variable in these procedures, this sequence provided a partial counterbalancing of the positions of the forced-choice terminal-link keys. After another 74 sessions of other procedures (cf. Fig. 18), Sequence D examined preference by maintaining free choice first in left terminal links and then in right terminal links while varying the number of available free-choice keys. Finally, after 17 additional sessions of other procedures, Sequence E again examined preference for two versus four lit keys in terminal links.

Free-choice preference as a function of the number of alternatives is summarized in Fig. 13; the letters correspond to the sequences of Fig. 12. Although the monotonic increase in preference with number of free-choice alternatives for Pigeon 3 in Sequence D is visually prominent, comparisons across pigeons and sequences do not demonstrate any consistent effect of number of alternatives on magnitude of free-choice preference. Of the nine comparisons between two and three free-choice alternatives (2:1 and 3:1), preference for three alternatives was greater than that for two in only four cases; of the nine comparisons between three and four free-choice alternatives (3:1 and 4:1), preference for four alternatives was greater than that for three in only four cases. Averaged either across sequences or across pigeons, preference was roughly constant at about .06 over the range from two to four free-choice alternatives. Preference for four free-choice alternatives over two (4:2), however, was typically smaller than that for two alternatives over one (2:1); the exception was Pigeon 18 in Sequences D and E. Averaged across sequences and pigeons, the 4:2 preference was about .03, or about half the 2:1 free-choice versus forced-choice preference. Over the present range, there is no evidence that free-choice preference decreases with number of free-choice alternatives, as suggested by the concept of overchoice (Toffler, 1970).

A weak intransitivity among free-choice preferences is implicit in these

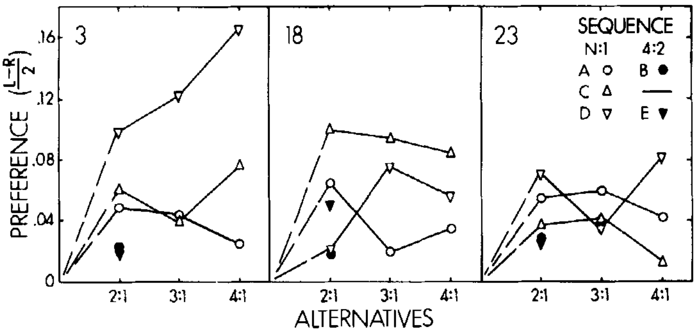


Fig. 13. Magnitude of preference as a function of number of alternatives for each of three pigeons. For each symmetrical pair of conditions (e.g., two keys versus one key and one key versus two keys), magnitude of preference was calculated as one-half the difference between relative initial-link rate with the larger number of alternatives in the left terminal link and that with the larger number of alternatives in the right terminal link. Each data point is the arithmetic mean across all relevant comparisons within each lettered sequence in Fig. 12.

data. On the one hand, it might be argued that, because both involve the same ratio between terminal-link alternatives, the 4:2 preference should be equal to a 2:1 preference; on the other, the roughly equal 2:1 and 4:1 preferences are consistent with indifference between the terminal links of the 4:2 condition. In any case, the independence of the magnitude of free-choice preference from the number of free-choice alternatives poses problems for various interpretations of the preference. For example, accounts in terms of the number of conditioned reinforcers, summation of the response probabilities occasioned by each alternative, or the opportunities for more variable responding provided by the larger total key area in free choice seem all to imply that the magnitude of free-choice preference should increase monotonically with number of alternatives and that the 4:2 preference should equal the 2:1 preference. To accommodate the present findings, therefore, such accounts would have to assume a nonmonotonic relation between preference and the relevant dimension of behavior (e.g., the function relating preference to available key area would have to pass through a maximum at some optimal key area in the neighborhood of 10 to 15 cm^2, i.e., the approximate area of two or three standard pigeon keys). Because a quantitative treatment of such accounts would be based solely on the preferences obtained in the present studies and would necessarily involve a number of ad hoc assumptions, no attempt will be made here to elaborate further on these alternative interpretations.

PREFERENCE AND CONCURRENT-CHAIN PARAMETERS

The present research was initially concerned simply with demonstrating a direction of preference. But relative magnitudes of preference are critical to interpretations of free-choice alternatives and of the relation between free choice and informative stimuli in terminal links. One difficulty is the deviation of baseline preference from .5 across pigeons and successive procedures. A .1 shift in preference from a baseline level of .4 cannot be assumed equivalent to a .1 shift from a baseline level of .5. But other problems are more serious. The quantitative analysis of preferences within concurrent chains remains controversial because preferences may vary with parameters of these schedules and because the ordering among preferences is not necessarily transitive (e.g., Navarick & Fantino, 1972).

For example, in a study of the effects of mixed versus multiple schedules in the terminal links of concurrent chains, preference reversed with changes in the initial-link VI schedules (Daniels, 1971). Concurrent and equal initial links were varied from VI 0-sec to VI 120-sec schedules;

mutually exclusive terminal links arranged food reinforcers according to mixed FI 10-sec FI 40-sec schedules alternating irregularly on a yellow key or according to multiple FI 10-sec FI 40-sec schedules alternating irregularly on respective green and red keys. At VI 0-sec, the higher rate of initial-link responding was on the key that produced the multiple-schedule terminal link; at VI 15-sec, initial-link response rates were roughly equal; and from VI 30-sec to VI 120-sec, the higher rate of initial-link responding was on the key that produced the mixed-schedule terminal link. A complication is the difficulty of equating number of exposures to the different terminal links as the VI schedule in the initial links becomes short. Nevertheless, conclusions drawn from concurrent-chain performances must be tempered by the possible dependence of preference on schedule parameters.

Some effects of concurrent-chain schedules may also depend on the relation between initial-link and terminal-link parameters. For example, consider a concurrent chain with an FI 30-sec schedule in the left terminal link and an FI 90-sec schedule in the right terminal link. If concurrent VI 120-sec schedules operate in the initial links, then the organism can produce about one reinforcer every 150 sec by responding exclusively in the left chain (120-sec initial link plus 30-sec terminal link) and about one every 210 sec by responding exclusively in the right chain (120-sec initial link plus 90-sec terminal link). If it distributes its responding to both chains, it will produce about one reinforcer every 120 sec (an average of 60 initial-link sec between successive terminal links, with terminal links averaging 60 sec). In other words, the contingencies favor the distribution of responses to both concurrent chains, because that performance produces reinforcers more often than exclusive responding in either chain alone. But the contingencies are different if concurrent VI 30-sec schedules are substituted in the initial links. Now the organism can produce about one reinforcer every 60 sec by responding exclusively in the left chain (30-sec initial link plus 30-sec terminal link) and about one every 120 sec by responding exclusively in the right chain (30-sec initial link plus 90-sec terminal link). The former rate of reinforcement is higher than that of about one every 75 sec produced by distributing responses to both chains (an average of 15 initial-link sec between successive terminal links, with terminal links averaging 60 sec). In this instance, the contingencies favor exclusive responding in the left chain. If some other variable were introduced into the terminal links of these concurrent chains, the effects of that variable on preference might be evident with VI 120-sec initial-link schedules but might be masked by the contingencies introduced with VI 30-sec initial-link schedules. Thus, some intransitivities may come about because terminal-link variables are effective at some parameter values and not others. The present concurrent chains, however, with VI initial links and equal FI terminal links, fall outside the range of schedule parameters and involve different initial-link and terminal-link schedule combinations than those for which intransitivities have been demonstrated (cf. Navarick & Fantino, 1972).

Terminal-link schedule parameters may affect preference even when the schedules are equal. In the present concurrent chains, the magnitude of free-choice preference increased with increases in the equal terminal-link FI schedules, as illustrated in Fig. 14. These data were obtained during the preliminary training that preceded the sessions of Fig. 11 (Catania & Sagvolden, 1980, Fig. 2). The rationale for lengthening the FI terminal links was based on the observation that the stimuli correlated with free choice or forced choice were presented at the onset of a given terminal link whereas the food reinforcer was delivered at the end of the fixed interval. Thus, the effectiveness of food as a reinforcer should increase with its immediacy, I (the reciprocal of the FI), whereas introducing free choice into a terminal link should add a constant reinforcing effect, C, to that terminal link. If preference depends on the reinforcers in a given terminal link relative to total reinforcers, then the relative reinforcing effects of free-choice and forced-choice terminal links are respectively functions of the following expressions: $(I + C)/[(I + C) + I]$ and $I/[(I + C) + I]$. The difference between these two expressions approaches zero as I becomes large (short FI, approaching immediate reinforcement) and approaches 1.0 as I decreases (lengthened FI). A complication, how-

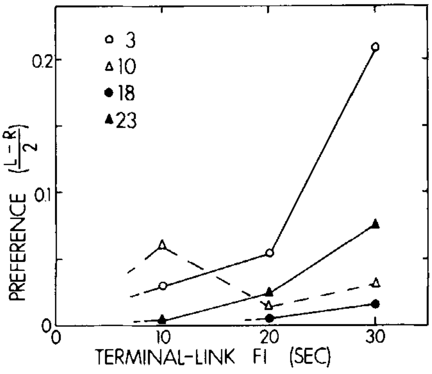


Fig. 14. Preference as a function of the duration of the terminal-link fixed-interval (FI) schedule, for four pigeons. Each data point shows one-half the difference in preference between concurrent chains with free-choice terminal links on the left and those with free-choice terminal links on the right. Pigeon 18 was not exposed to both conditions with FI 10-sec terminal links; the data for Pigeon 10 are connected by dashed lines because the absolute initial-link response rates were low and variable for this pigeon throughout all conditions. (Adapted from Catania & Sagvolden, 1980, Fig. 2.)

ever, is that lengthening the terminal-link FI may weaken overall initial-link responding at the same time it enhances free-choice preference. In any case, the generality of free-choice preference depends on its demonstration in a variety of procedures. Although the magnitude of free-choice preference varies with schedule parameters, this preference was neither absent nor reversed over the range of conditions examined here.

V. Free Choice among Different Reinforcers

> The fault was Nature's fault not thine,
> Which made thee fickle as thou art.
> LORD BYRON

> Doth not the appetite alter?
> WILLIAM SHAKESPEARE

The research thus far has examined preference for free choice among different routes to the same reinforcer. The present section extends the account to preference for free choice among different reinforcers. These experiments explicitly address a property of terminal-link performance that was only implicit in the earlier procedures: the role of momentary changes in preference from one terminal link to the next. When a free choice between two different reinforcers is available, the reinforcer that the organism produces is by definition the one that is momentarily preferred, but this preference may change over later opportunities for free choice. When a forced choice of one of the two reinforcers is arranged, however, the reinforcer that the experimenter has made available may or may not be the one that is momentarily preferred (cf. Brigham & Sherman, 1973). Thus, a preference for free choice over forced choice may come about simply because the momentarily preferred reinforcer is always available in free-choice terminal links whereas it is only sometimes available in forced-choice terminal links.

Concurrent-chain schedules were arranged in which food and water reinforcers were available in terminal links. Methods and apparatus for arranging these reinforcers are well-established, and more data are available on performances maintained by food and water reinforcers than on those maintained by other combinations of reinforcers (e.g., Hursh, 1978; Willis, Van Hartesveldt, Loken, & Hall, 1974). These factors outweigh the disadvantage that food deprivation and water deprivation have interacting effects: food deprivation reduces the effectiveness of water as a reinforcer, and water deprivation reduces the effectiveness of food as a reinforcer (Zeigler, Green, & Seigel, 1972). For any pair of different reinforcers, it would be inappropriate in any case to assume that schedule interactions would be comparable to those when reinforcers are the same (e.g., Steinman, 1968).

The apparatus is illustrated in Fig. 15. Concurrent and independent VI 30-sec initial links operated on the two center keys, lit white (W), with the side keys dark. The first peck after a changeover from one initial-link key to the other was not eligible to produce a terminal link (changeover ratio 2). Pecks on one initial-link key produced free-choice terminal links; pecks on the other produced forced-choice terminal links. During terminal links, the center keys were dark and one or both side keys lit. In free-choice terminal links, the left key was amber (A) and the right key was blue (B). At the end of a fixed interval, the first peck on either key was reinforced: an amber-key peck produced a 5-sec food delivery in the standard Gerbrands pigeon feeder on the left, and a blue-key peck produced a 5-sec delivery of a .2-cm^3 water cup in a standard Gerbrands dipper feeder on the right. In forced-choice terminal links, either the left key was lit amber, making a food reinforcer available at the end of the fixed interval, or the right key was lit blue, making a water reinforcer available. The terminal-link schedule was FI 15-sec through the first 175

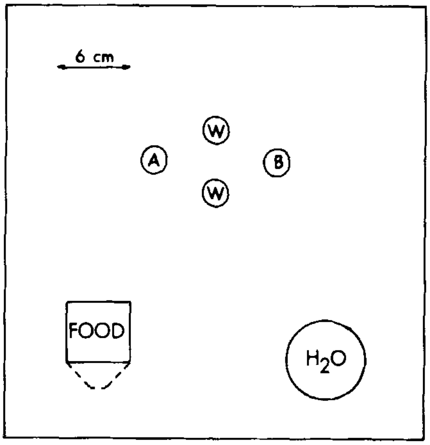


Fig. 15. Experimental panel for the study of concurrent-chain schedules with different reinforcers arranged in terminal links. Independent concurrent variable-interval (VI) initial links were arranged on the two middle white (W) keys. Pecks on one of the initial-link keys produced forced-choice terminal links: either the left key lit amber (A) and a peck on this key produced food at the end of a fixed interval (FI), or the right key lit blue (B) and a peck on this key produced water at the end of the FI. Pecks on the other initial-link key produced free-choice terminal links: both side keys lit, and food or water was presented depending on which key was pecked at the end of the FI. When food or water was delivered, the relevant device was lit, other chamber lights were turned off, and eating and drinking were monitored with photocells.

sessions (condition I in Fig. 16), and FI 20-sec thereafter. During reinforcer deliveries, the relevant feeder was lit and the other lights were off; after the delivery of either reinforcer, the initial links were reinstated.

The daily volume of water intake of three 1-year-old male Silver King pigeons, maintained at 80% of free-feeding body weight, was measured for 18 days. Levels of food and water deprivation were then maintained by postsession feeding and by a postsession water ration initially equal to the difference between the preceding mean daily water intake and the estimated water intake within sessions (visual observation and photocell monitoring indicated that the pigeons rarely produced a food or water reinforcer without consuming it). The daily water ration was occasionally adjusted upward or downward on the basis of performance, and session durations were also adjusted to minimize the postsession rations. With these arrangements, each pigeon typically obtained all of its daily food but only a fraction of its daily water within experimental sessions. At these levels, each pigeon's pecking was maintained on both the amber and the blue key when either was presented alone in a forced-choice terminal link; the distribution of responses to the two keys when they were presented together in free-choice terminal links varied from pigeon to pigeon. Preliminary research had shown that increases in water deprivation beyond these levels reduced food maintained responding to such an extent that the pigeons could not be maintained at 80% of free-feeding weights.

The range of session durations was 30 to 45 min of initial links for Pigeon 37, 15 to 30 min for Pigeon 43, and 45 to 60 min for Pigeon 40. Experimental conditions in which the free-choice terminal link was produced by the top initial-link key alternated with those in which it was produced by the bottom initial-link key. Conditions were typically maintained for 7 to 14 sessions, but some conditions were extended because of variable performance that presumably was attributable to problems in the maintenance of deprivation levels. Pigeons were moved from one condition to the next individually rather than as a group; thus, the number of sessions at a given condition varied across pigeons.

Figure 16 summarizes the data obtained with food and water reinforcers available in the free-choice terminal link of these concurrent-chain schedules. The present initial links operated on top (T) and bottom (B) keys, but for consistency with previous figures the x-axis is again arranged so that respective left and right displacements of apex-left and apex-right triangles correspond to shifts of preference toward free-choice terminal links. In conditions A through M, only the amber key (food) was available in forced-choice terminal links; in the remaining conditions, only the blue key (water) was available. Baseline conditions (equal termi-

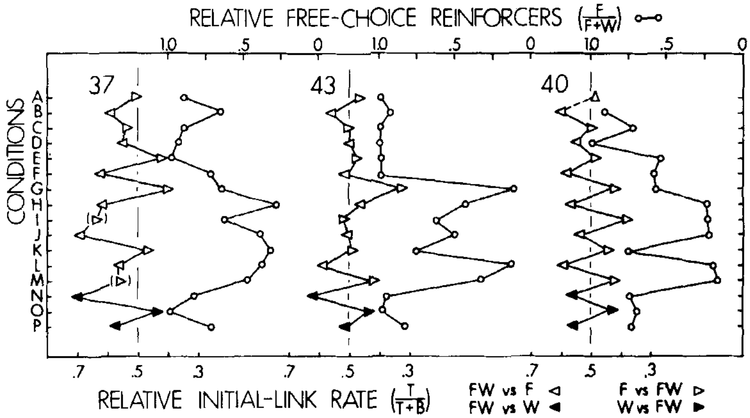


Fig. 16. Relative initial-link response rates and relative free-choice reinforcers with food (F) and water (W) terminal-link reinforcers for each of three pigeons. Successive conditions are arranged along the *y*-axis. For consistency with previous figures, relative initial-link rates are shown along the *x*-axis, with displacements to the left corresponding to increased top-key preference (cf. Fig. 14). Apex-left triangles indicate free choice in top terminal links; apex-right triangles indicate free choice in bottom terminal links. The open circles, plotted against the upper scale, show food reinforcers as a proportion of total free-choice reinforcers in each condition. Each data point is an arithmetic mean across the last three sessions of a condition.

nal links) were maintained for Pigeon 40 during the first experimental conditions (A) for Pigeons 37 and 43. The bracketed data points indicate low and variable initial-link response rates for Pigeon 37 during conditions I and M.

The triangles show preference (relative initial-link responding) over successive conditions. The circles show the proportion of food reinforcers produced in the free-choice terminal link (relative free-choice reinforcers) in those conditions. In general, free choice was preferred to forced choice, but except for Pigeon 40 the preference was not obtained as reliably as in the preceding studies. As in the earlier experiments, preference was not correlated with relative terminal-link response rates. In conditions A through M, with only food in forced-choice terminal links, the pigeons typically produced both food and water in free-choice terminal links. The exceptions were conditions A, C, D, E, and F for Pigeon 43 and condition D for Pigeon 40; in condition E food constituted all but 3% of the free-choice reinforcers for Pigeon 37. In conditions N through P, with only water in forced-choice terminal links, the proportion of food reinforcers produced in free-choice terminal links increased over that during conditions L and M for all pigeons.

In conditions A through M, water was sometimes the momentarily preferred reinforcer but was available only in free-choice terminal links.

Similarily, in conditions N through P, food was sometimes the momentarily preferred reinforcer but was available only in free-choice terminal links. Thus, the difference in the availability of reinforcers in free-choice and forced-choice terminal links provides a basis for the observed free-choice preferences. In subsequent conditions, an attempt was made to assess the magnitude of free-choice preference relative to preference for food versus water reinforcers in terminal links; the results were equivocal because of problems with equating deprivation conditions across different experimental procedures.

THE PROBLEM OF MOMENTARY CHANGES IN REINFORCER PREFERENCE

The problem of momentary changes in reinforcer preference can be resolved only if, when a forced-choice terminal link is presented, the experimenter can determine which reinforcer the organism would have produced if a free-choice terminal link had been presented instead. One circumstance in which the experimenter can do this is when the organism exclusively produces only one of the two reinforcers available in free-choice terminal links. That reinforcer can then be the one made available in forced-choice terminal links. At the beginning of the experiments of Fig. 16, the performance of Pigeon 43 satisfied this criterion. During the baseline sessions that preceded condition A, the concurrent chains were arranged with equal terminal links. Three conditions, forced choice of food in both terminal links, forced choice of water in both terminal links, and free choice of food or water in both terminal links, alternated irregularly. Throughout these sessions, Pigeon 43 produced water only in terminal links that arranged forced choice of water; with both forced choice of food and free choice of food or water this pigeon pecked exclusively on the food terminal-link key (amber). During these baseline sessions, with equal terminal links, there was a substantial preference for the top initial-link key.

Relative initial-link response rates over the first 25 experimental sessions for Pigeon 43 are shown in Fig. 17. The arrangement of the x-axis and of the apex-left and apex-right triangles is as in previous figures. When forced choice of food was arranged in the terminal link produced by top-key initial-link pecks and free choice of food or water was arranged in the terminal link produced by bottom-key initial-link pecks (sessions 1 to 7), preference shifted to the bottom initial link (to the right in Fig. 17). When the terminal links were reversed (sessions 8 to 15), preference shifted back toward the top initial link (left in Fig. 17). With a return to the former condition (sessions 16 to 25), preference again shifted

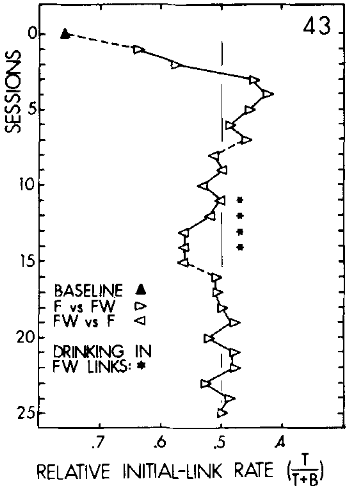


Fig. 17. Relative initial-link response rates over successive sessions for Pigeon 43. In all sessions except those marked with asterisks, this pigeon pecked only the food key in food-water free-choice terminal links. The data presentation corresponds to that of Fig. 15.

toward the bottom initial link, though to a lesser extent. In sessions 11 through 14 (asterisks), Pigeon 43 produced water in some free-choice terminal links (three water reinforcers in session 11, five each in sessions 12 and 13, and one in session 14); during these sessions (April 1976), the laboratory air conditioners were not operating during an unanticipated heat wave and the laboratory temperature rose to more than 32°C (90°F). Drinking in free-choice terminal links stopped when the temperature returned to normal levels and did not begin again until condition G, but the free-choice preference diminished with successive reversals (cf. Fig. 16).

Although a free-choice preference was obtained, these data for Pigeon 43 are only suggestive. A conclusion about preference for free choice among different reinforcers must await further research. One possibility is to deliver food and water reinforcers in only some fraction of the free-choice terminal links, and to use the performances in the remaining free-choice terminal links as probes of momentary changes in reinforcer preference upon which the succession of food and water forced choices in the other terminal link could be based. Another is to arrange concurrent chains with three initial-link keys. One VI schedule would operate for pecks on the first initial-link key, which would produce terminal links with free choice of food or water. A second and equal VI schedule would operate for pecks on either of the remaining two initial-link keys; these pecks would each produce forced-choice terminal links, but those on one key would produce forced choice of food whereas those on the other would produce forced choice of water. Presumably of the two latter initial-link keys the pigeon would peck the one that produced forced choice of the momentarily preferred reinforcer.

VI. The Magnitude and Durability of Free-Choice Preference

> . . . choice can be manipulated as readily as it can be coerced.
> P. H. PARTRIDGE

> . . . the reaction against freedom is almost everywhere triumphant. . . . The people have been promised abundance, security, peace, if they would surrender the heritage of liberty.
> WALTER LIPPMAN

> A hungry man is not a free man.
> ADLAI STEVENSON

Table III summarizes the magnitudes of preference obtained in the various concurrent-chain studies. With two, three, or four free-choice alternatives in terminal links, preference for free choice over forced choice was about .06; preference for four over two free-choice alternatives was about .03, or half the free-choice preference. Although preferences were demonstrated for two lit keys over one lit key and for varied key colors over unvaried key colors, these preferences of about .02 were small relative to those for free choice over forced choice. On the other

TABLE III

MAGNITUDE OF PREFERENCE FOR VARIOUS CONCURRENT-CHAIN TERMINAL LINKS, EXPRESSED AS DISPLACEMENT FROM BASELINE (ONE-HALF THE ABSOLUTE DIFFERENCE BETWEEN X-ON-LEFT AND X-ON-RIGHT)[a]

Comparison	Terminal link X	Terminal link Y	Magnitude of preference
1. 2:1	2 FI keys	1 FI key	.06
2. 3:1	3 FI keys	1 FI key	.06
3. 4:1	4 FI keys	1 FI key	.06
4. 4:2	4 FI keys	2 FI keys	.03
5. Lit keys	1 FI key, 1 EXT key	1 FI key	.02
6. Variety	2 different-color keys	2 same-color keys	.02
7. Information	1 FI key, one different-color EXT key	1 FI key, one same-color EXT key	.09

[a] Data for comparisons 1 through 4 are from Fig. 13; Fig. 4 provides the same value for comparison 1 and Fig. 11 provides the same value for comparison 2. Data for comparisons 5 through 7 are from Table 3 in Catania (1975); see also the present Fig. 9.

hand, the preference of .09 for informative stimuli over uninformative stimuli was greater than that for free choice over forced choice. The demonstration of an important role of informative stimuli is appropriate to an analysis of free-choice preference, because informative stimuli are an integral component of the availability of free choice. The availability of alternative responses can have little behavioral significance unless it is correlated with discriminative stimuli. The discriminative stimuli are not simply opportunities for responding; they are the occasions on which those responses produce particular consequences. This is the sense in which freedom and knowledge are inextricably related (as when we speak of being both free and informed; cf. Catania, 1975).

The preferences in Table III are expressed as estimates of displacements from baseline. For example, if relative responding in the left initial link is .58 with free choice in the left terminal link and changes to .46 when free choice is moved to the right terminal link, the estimated baseline is .52 and the magnitude of preference is .06. If the overall rate of food reinforcement over the initial and terminal links of the concurrent chains were 100 rf/hr (reinforcers per hour), then, on the assumption of approximate matching of relative response rate to relative reinforcement rate (e.g., Herrnstein, 1964), this shift in preference would be equivalent to that produced by moving 12 rf/hr from one concurrent chain to the other (respective left and right reinforcement rates of 56 and 44 rf/hr changed to rates of 44 and 56 rf/hr). Thus, the effect of preference of such variables as free choice and informative stimuli can be expressed in terms of equivalent reinforcement rates.

Nevertheless, the dimensional status of these variables is yet to be determined. Number of alternatives does not seem to be an appropriate metric for free choice, because magnitude of preference does not vary with this parameter. Similarly, the traditional information measure, bits, does not seem to be correlated with the behavioral effects of informative stimuli (e.g., Gollub, 1970), although that statistic may be appropriate for the variable here called variety (Fig. 9). Even the presumably unambiguous variable of number, in the preference for two lit keys over one lit key, may on analysis be replaceable by key area or total luminance of the keys. In any case, number and rate are also not dimensionally compatible.

The interpretation of the research on informative stimuli (Fig. 9) was that in some circumstances preference for informative stimuli in one terminal link could counterbalance preference for free choice in the other terminal link. In general, if the availability of free choice can function as a reinforcer in one terminal link, it also follows that the free-choice preference can be overridden by the reinforcers available in the other terminal link. This outcome is illustrated in Fig. 18 (session numbers

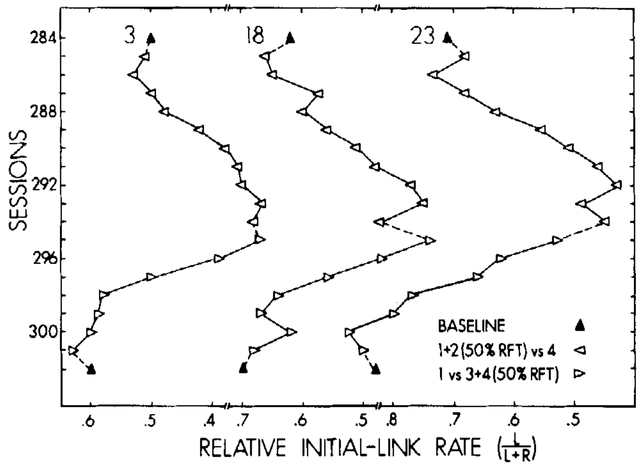


Fig. 18. Relative initial-link response rate over successive sessions for three pigeons. The procedures correspond to those of Fig. 12, but half of the reinforcers in free-choice terminal links were omitted.

correspond to those along the right y-axis of Fig. 12). The concurrent-chain stimuli and schedules were as in Fig. 12, except that pecks were reinforced with food at the end of only half the free-choice terminal links; in the other half, irregularly alternating, the key lights were out for 3.5 sec but the feeder was not operated. Pecks produced food at the end of the fixed interval in every forced-choice terminal link.

After baseline schedules with forced choice in both terminal links, free choice of Keys 1 and 2 (red) was arranged in left terminal links and forced choice of Key 4 (green) was arranged in right terminal links. Over 10 daily sessions, preference shifted toward the forced-choice right terminal links, in which all reinforcers were presented, and away from the free-choice left terminal links, in which half the reinforcers were omitted. The terminal links were then reversed, with forced choice of Key 1 (red) arranged in left terminal links and free choice of Keys 3 and 4 (green) arranged in right terminal links. Over the next seven sessions, preference again shifted away from free-choice toward forced-choice terminal links, in the direction corresponding to 100% rather than 50% food reinforcement. Based on means across the last three sessions of each condition, the shift in preference was from a left initial-link relative response rate of about .41 to one of about .70 (cf. the respective values of .33 and .67 that would be predicted from matching of relative response rates to relative reinforcement rates). Expressed as a displacement from baseline, as in Table III, this is equivalent to a preference of about .15 for forced-choice 100% reinforcement over free-choice 50% reinforcement.

The avoidance of free choice has been discussed in the literature of freedom (e.g., Fromm, 1941). Such accounts are consistent with the demonstration that other reinforcement variables can override the free-choice preference. In addition, research may be in order on whether free choice is preferred to forced choice when the alternative responses are maintained by avoidance or escape contingencies, or when the alternative responses are not only maintained by reinforcers but also produce punishers.

In studies of escape from conflict, a forced-choice preference is also implicit. For example, in experiments on self-control in pigeons (e.g., Rachlin & Green, 1972) one response produces a small immediate reinforcer and another produces a large but delayed one. The pigeon typically produces the small immediate reinforcer rather than the larger delayed one. In the time preceding the opportunity to produce these reinforcers, however, a commitment response may be made available. If the commitment response occurs, it cancels the later availability of the small immediate reinforcer and thereby guarantees that the pigeon will produce the larger delayed one. In other words, the commitment response changes the later situation from one of free choice between the two reinforcers to one of forced choice of the larger reinforcer. But maintenance of the commitment response in self-control procedures is not a failure of free-choice preference; rather, any free-choice preference in these situations is counterbalanced by the greater magnitude of reinforcement when commitment responses limit the pigeon to forced choice of the larger but delayed reinforcer.

VII. Ontogenic and Phylogenic Sources of Free-Choice Preference

> . . . democracy gives the most freedom of choice to beings whose most urgent need is freedom of choice.
>
> John Fowles

> We no longer have the luxury of choice.
>
> James R. Schlesinger

The potent effects of alternatives or options have been demonstrated in the context of complex reinforcement schedules (Findley, 1962); some of those schedules maintained response outputs orders of magnitude larger than those maintained by standard reinforcement schedules. The significance of free-choice preference also seems implicit in applied behavior research, as when token systems arrange a choice of reinforcers (e.g., Brigham & Sherman, 1973; Brigham & Stoerzinger, 1976), or when students' self-imposed contingencies are more effective than teacher-imposed contingencies in modifying academic behavior (e.g., Lovitt & Curtis, 1969; Taffel & O'Leary, 1976). The concept of locus of control is closely related to the availability of free choice (e.g., Lefcourt, 1966, 1973), and this variable has also been implicated as behaviorally significant in studies of verbal learning (e.g., Monty, Rosenberger, & Perlmuter, 1973; Perlmuter, Monty, & Kimble, 1971), behavior therapy (MacDonough, Adams, & Tesser, 1973), and social psychology (Brehm, 1966; Langer & Rodin, 1976; Wicklund, 1974).

Because free choice implies the availability of alternatives, the concept of freedom occasions empirical questions. To argue that freedom should be valued, for example, is to argue that free choices should be preferred to forced choices. The present studies have demonstrated such preferences, and have shown that they are not limited to the behavior of humans. Such a finding does not demean free-choice preference; in fact, it makes the preference more fundamental, because if the preference exists even in the behavior of pigeons or rats it is not reducible simply to the product of some particular human culture. It may then be asked whether free-choice preference is a product of ontogeny or phylogeny. Does it exist because organisms learn that momentarily preferred alternatives are more often available in free-choice than in forced-choice conditions, or because evolutionary contingencies have favored the survival of organisms that prefer free choice to forced choice?

A. BEHAVIORAL CONTINGENCIES AND FREE-CHOICE PREFERENCE

The independence of the magnitude of free-choice preference from the number of free-choice alternatives reduces the plausibility of ontogenic accounts that appeal to conditioned reinforcement or other acquired behavioral properties of stimuli. Nevertheless, it is useful to consider how interactions with the environment during an organism's lifetime can lead to such preferences.

If a pigeon's pecks on one key are ineffective and another key is available, switching to and pecking the other key may be reinforced. Such switching, and subsequent reinforcement, is not possible when no other key is available. Exposed to such contingencies, the pigeon may learn that one-key situations sometimes lead to periods of nonreinforcement whereas more-than-one-key situations do not. In the present experiments, such contingencies could have been produced by failures of terminal-link keys. Terminal links ended only with the delivery of a reinforcer. Thus, in the event of a key failure in a free-choice terminal link the pigeon could terminate the fixed interval by switching to another key. But in a forced-choice terminal link the pigeon remained in the presence of the single inoperable key because a peck on that key was ineffective in producing the reinforcer at the end of the fixed interval. Unless the key failure was intermittent, the situation did not change until the experimenter intervened.

In the present research, however, there was no evidence that magnitude of free-choice preference changed systematically with apparatus failures. In the first experiments (Fig. 4), magnitude of free-choice preference did not vary systematically with continued exposure to the concurrent-chain procedure, and was smallest for the pigeon (280) that had the most frequent exposure to terminal-link key failures. In the later experiments, key failures, detected on the basis of latencies from the end of the terminal-link FI to the reinforced peck, were infrequent. Of the 108 sessions of concurrent-chain procedures that preceded the sessions of Fig. 11, the only terminal-link key failure occurred in a single session for Pigeon 10; the other three pigeons entered the procedures of Fig. 11 without any known history of exposure to apparatus failures. During the 132 sessions of Fig. 11, terminal-link key failures were detected and corrected on four occasions for Pigeon 3, two for Pigeon 10, two for Pigeon 18, and three for Pigeon 23. These occasions were not followed by any systematic changes in the magnitude of the free-choice preference. Even the temporary correlation of free choice with reduced reinforcement rates did not affect the subsequent magnitude of free-choice preference (cf. Figs. 12 and 18). By describing how free choice but not forced choice can leave the organism with another response to fall back on when contingencies change, the preceding ontogenic argument could conceivably be extended to situations outside of the experimental environment. But such arguments, however plausible, are not likely to be convincing in the absence of evidence that free-choice preference varies as a function of exposure to free-choice and forced-choice contingencies.

Another behavioral property of free-choice contingencies was treated most explicitly in connection with free choice among different reinforcers (Section V), but is relevant to any free-choice situation. Whether free choice involves different reinforcers or different routes to the same reinforcer, momentary preferences among the alternatives may change over successive free-choice and forced-choice opportunities. The point is illustrated with respect to different terminal-link key colors and positions in Fig. 19. On the two keys of a free-choice terminal link, red (R) is available on the left and green (G) on the right. On the basis of prior sequences of choices, a sequence of left red and right green keys is made

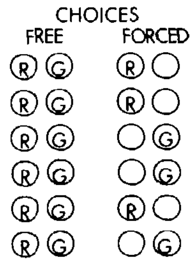


Fig. 19. A sequence of free choices and forced choices. Arranging the sequence of alternatives available in successive forced-choice conditions so that it corresponds to the sequence chosen in successive free-choice conditions does not guarantee that the momentarily preferred alternative will be presented in every forced-choice condition.

available in corresponding forced-choice terminal links. No matter how indifferent the pigeon is to key color and location, if there is some variability in the pigeon's choices, statistical predictions will occasionally present the key in a forced-choice terminal link that would not have been pecked if the terminal link had instead provided a free choice. In the forced-choice sequence shown, for example, the key colors presented were RRGGRG; the third of these was the momentarily less preferred alternative if the pigeon would have pecked the keys in the sequence RRRGRG had these been free-choice rather than forced-choice terminal links.

In the discussion of free choice among different reinforcers, this aspect of the concurrent-chain performances was treated as a feature to be controlled. To the extent that the reinforcers delivered in free-choice terminal links were not equivalent to those delivered in forced-choice terminal links, the magnitude of the free-choice preference could not be assessed independently of the difference in reinforcers. There is a sense, however, in which the difference between free-choice and forced-choice reinforcers should be regarded not as an incidental procedural artifact but rather as the behavioral basis of free-choice preference. If preferences among alternatives change from moment to moment, a free-choice preference may develop as the organism learns that momentarily preferred alternatives are always available in free-choice terminal links but are occasionally unavailable in forced-choice terminal links. A preference based on such learning might be manifested even over sessions in which free-choice and forced-choice reinforcers were perfectly matched. It is reasonable to speculate that such learning, based on a history of free choices and forced choices, might generalize across a variety of experimental and extraexperimental settings.

B. EVOLUTIONARY CONTINGENCIES AND FREE-CHOICE PREFERENCE

Whatever the status of the arguments for a behavioral origin of free-choice preference, the possibility of a biological origin also must not be excluded (''Phylogeny and ontogeny are friendly rivals and neither one always wins'': Skinner, 1977, p. 1009; see also Skinner, 1966, 1975). Organisms that prefer the availability of different responses or opportunities for more variable behavior may have evolutionary advantages (e.g., Hogan, 1971, 1973; Kavanau, 1969, p. 268; Roeder, 1975; Rozin & Kalat, 1971). For example, given that food supplies sometimes may be lost to competitors or may disappear in other ways, an organism that chooses environments in which two or more food supplies are available will probably have a survival advantage over one that chooses environments containing only a single food supply.

It is easy to multiply examples of how the evolutionary contingencies produced by forgiving and unforgiving environments might favor free-choice preference. The arctic mammal that prefers to dive for fish in an ice-covered pool with two breathing holes may be more likely to survive if the temperature drops and the breathing holes begin to freeze over than one that prefers to dive in another pool with only a single breathing hole. The nesting animal that prefers sites with alternate escape routes may be less likely to be caught by a predator than one that prefers sites with a single escape route. The nomadic animal that prefers to range over an area that includes several watering holes may be more likely to survive a drought than one that prefers an area with a single watering hole. But the issue will not be resolved merely by the production of plausible examples.

Variability provides the raw material upon which both ontogenic and phylogenic contingencies operate. The shaping of new classes of responses depends on the range of responding available for differential reinforcement, just as the evolution of species depends on variation among the members of populations. If two populations of either responses or organisms differed only in variability, the population with the greater variability would be more likely to survive in the face of changing contingencies. To the extent that free choice provides an opportunity for more variable responding than does forced choice, the free-choice preference is at least consistent with the important behavioral and biological role of variability.

A phylogenic basis for free-choice preference does not imply that the preference cannot be overridden by other variables. Figure 18 showed that a forced-choice preference could be established by reducing free-choice but not forced-choice reinforcers. A similar result could presumably be generated by punishing free-choice but not forced-choice responses. If free-choice preference has biological roots, however, the implication is that such procedures cannot eliminate it but can only mask it. If so, the ontogenic contingencies generated by such masking procedures are likely to be short-lived relative to the phylogenic contingencies that created the free-choice preference in the first place; the free-choice preference is then an abiding one even if it is sometimes relatively small in magnitude.

C. FREE CHOICE, BEHAVIOR ANALYSIS, AND BEHAVIOR SYNTHESIS

The empirical status of free-choice preference is orthogonal to the issue of free will and determinism. An account of the circumstances under which organisms produce an opportunity for free choice leaves open the philosophical question of whether they can then choose freely or must choose in predetermined ways. The definition of free choice, however, necessarily interacts with experimental findings. Once a free-choice preference has been demonstrated, the definition can be refined on the basis of research that explores the boundary conditions of the preference. For example, the possibility that dark and inoperative keys should be counted as free choices was initially ruled out on the grounds that every procedure in the six-key apparatus would count as a free-choice condition, and extinction keys either lit or dark were then rejected as potential free-choice alternatives on the basis of experimental evidence (e.g., Figs. 10 and 11). The finding that the magnitude of preference was independent of the number of free-choice alternatives (Fig. 13) argued against free choice as a derivative of the key area available for reinforced pecking, and the equivalence of characterizing free choice either as the availability of discriminative stimuli or as the opportunity for varied reinforced responding was implicit in the analysis of the role of momentary changes in preference. Future research that restricts free choice by punishing some of the alternative responses may further delineate the limits of free choice by providing a behavioral basis for distinguishing between free choices and coerced choices. Other procedures might bear on the question of whether the opportunity for responding *versus* not responding constitutes a choice.

To speak of free choice in terms of discriminative stimuli or opportunities for varied responding or the availability of momentarily preferred reinforcers is to emphasize separately the several features that together make up the behavioral classes called discriminated operants. In the presence of a discriminative stimulus (S^D) a response (R) produces some reinforcing consequence (consequential stimulus, or S^C). Discriminated operants are the basis for a paradigm or symbolic representation of the present concurrent-chain procedures, shown in Fig. 20). A response in one initial link produces forced choice, a single discriminated operant (1). A response in the other initial link produces free choice, two discriminated operants (2 and 3). To the extent that the discriminative stimuli of the different discriminated operants maintain the respective initial-link responses, R_x and R_y, these stimuli are also conditioned reinforcers. But because some consequences (e.g., those of pecking unlit keys) are not adequate in generating conditions of free choice, the discriminated operants must be defined to exclude such inadequate cases.

One possibility for a behavioral definition of free-choice alternatives rests with the observation that each of the discriminated operants available together in free-choice terminal links maintained initial-link responding when it was made available separately in forced-choice terminal links. For example, in the experiment on free choice of food and water, Pigeon 43 produced only food but not water in early free-choice terminal links, but did produce water when it was available alone in forced-choice terminal links (cf. Section V and Fig. 17). In the paradigm of Fig. 20, therefore, the terminal-link responses R_2 and R_3 should each be maintained in the presence of their respective stimuli when they occur alone in forced choice, even if both are not maintained when their stimuli are presented together in free choice. The advantage of such a definition of the availability of free choice is that it provides a behavioral criterion that is independent of free-choice preference. The relation of this criterion to the preference then becomes an empirical issue rather than an issue of the vocabulary of free choice.

Behavior analysis begins with complex behavioral relations and breaks them down into their components. The basic processes that emerge from

$$R_x — \left[S^D_1 \cdot (R_1 — S^C_1)\right]$$

$$R_y — \left[\begin{array}{l} S^D_2 \cdot (R_2 — S^C_2) \\ S^D_3 \cdot (R_3 — S^C_3) \end{array}\right]$$

Fig. 20. Paradigms for forced choice (produced by response x) and free choice (produced by response y). The two conditions, shown in brackets, are represented by the number of available discriminated operants. A discriminated operant is defined by a discriminative stimulus (S^D) in the presence of which a response (R) produces a consequence (S^C). One problem in defining choice is that of specifying the kinds of consequences in free choice that will maintain the free-choice preference. Each consequence must be a reinforcer, but the preference must not depend on differences between S_1^C and S_2^C or S_1^C and S_3^C.

such analyses are sometimes seen to be relevant to important human concepts. But it is also possible to begin with the everyday language of human affairs as a basis for synthesizing important behavioral relations in the laboratory. The behavior synthesis then consists of producing in the laboratory a performance that is in some respects analogous to human behavior outside of the laboratory. Experiments on self-control (Rachlin & Green, 1972) and the present research on freedom of choice provide examples. The behavior synthesis would be of little interest, however, if it merely generated an analog; if that were its only function, a thought experiment would do as well. Instead, once a phenomenon has been demonstrated by a behavior synthesis, the definition of the phenomenon can then be refined through subsequent research. The success of the synthesis should therefore be judged not only on the basis of the empirical findings but also on the extent to which the refined definition of the phenomenon has implications for the nonlaboratory human situations from which the analog emerged.

REFERENCES

Autor, S. M. The strength of conditioned reinforcers as a function of frequency and probability of reinforcement. In D. P. Hendry (Ed.), *Conditioned reinforcement*. Homewood, Ill.: Dorsey, 1969.

Berlyne, D. E. Curiosity and exploration. *Science*, 1966, **153**, 25–33.

Berlyne, D. E. Reinforcement values of visual patterns compared through concurrent performances. *Journal of the Experimental Analysis of Behavior*, 1972, **18**, 281–285.

Berofsky, B. (Ed.) *Free will and determinism*. New York: Harper, 1966.

Boren, J. J., Moerschbaecher, J. M., & Whyte, A. A. Variability of response location on fixed-ratio and fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1978, **30**, 63–67.

Bower, G., McLean, J., & Meacham, J. Value of knowing when reinforcement is due. *Journal of Comparative and Physiological Psychology*, 1966, **62**, 184–192.

Brehm, J. W. *A theory of psychological reactance*. New York: Academic Press, 1966.

Brigham, T. A., & Sherman, J. A. Effects of choice and immediacy of reinforcement on single response and switching behavior of children. *Journal of the Experimental Analysis of Behavior*, 1973, **19**, 425–435.

Brigham, T. A., & Stoerzinger, A. An experimental analysis of children's preference for self-selected rewards. In T. A. Brigham, R. Hawkins, J. Scott, & T. F. McLaughlin (Eds.), *Behavior analysis in education: Self-control and reading*. Dubuque, Iowa: Kendall/Hunt, 1976.

Catania, A. C. The concept of the operant in the analysis of behavior. *Behaviorism*, 1973, **1**, 103–116.

Catania, A. C. Freedom and knowledge: An experimental analysis of preference in pigeons. *Journal of the Experimental Analysis of Behavior*, 1975, **24**, 89–106.

Catania, A. C., & Reynolds, G. S. A quantitative analysis of the behavior maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1968, **11**, 327–383.

Catania, A. C., & Sagvolden, T. Preference for free choice over forced choice in pigeons. *Journal of the Experimental Analysis of Behavior,* 1980, **34,** (in press).

Daniels, P. V. *The reinforcing value of information.* Unpublished doctoral dissertation, Columbia University, 1971.

Dashiell, J. F. A quantitative demonstration of animal drive. *Journal of Comparative Psychology,* 1925, **5,** 205-208.

Duncan, B., & Fantino, E. The psychological distance to reward. *Journal of the Experimental Analysis of Behavior,* 1972, **18,** 23-34.

Eckerman, D. A., & Lanson, R. N. Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement, and extinction. *Journal of the Experimental Analysis of Behavior,* 1969, **12,** 73-80.

Egger, M. D., & Miller, N. E. Secondary reinforcement in rats as a function of information value and reliability of the stimulus. *Journal of Experimental Psychology,* 1962, **64,** 97-104.

Enteman, W. F. (Ed.) *The problem of free will.* New York: Scribner, 1967.

Fantino, E. Effects of required rates of responding upon choice. *Journal of the Experimental Analysis of Behavior,* 1968, **11,** 15-22.

Findley, J. D. An experimental outline for building and exploring multi-operant behavior repertoires. *Journal of the Experimental Analysis of Behavior,* 1962, **5,** 113-166.

Fisher, M. A., & Catania, A. C. Autoshaping: Relation of feeder color to choice of key color. *Bulletin of the Psychonomic Society,* 1977, **9,** 439-442.

Fromm, E. *Escape from Freedom.* New York: Rinehart, 1941.

Gollub, L. R. Information on conditioned reinforcement. A review of *Conditioned reinforcement. Journal of the Experimental Analysis of Behavior,* 1970, **14,** 361-372.

Greenspan, P. S. Behavior control and freedom of action. *Philosophical Review,* 1978, **87,** 225-240.

Herrnstein, R. J. Stereotypy and intermittent reinforcement. *Science,* 1961, **133,** 2067-2069.

Herrnstein, R. J. Secondary reinforcement and rate of primary reinforcement. *Journal of the Experimental Analysis of Behavior,* 1964, **7,** 27-36.

Hobart, R. E. Free will as involving determination and inconceivable without it. *Mind,* 1934, **43,** 1-27.

Hogan, J. A. The development of a hunger system in young chicks. *Behaviour,* 1971, **39,** 128-201.

Hogan, J. A. How young chicks learn to recognize food. In R. A. Hinde & J. Stevenson-Hinde (Eds.), *Constraints on learning.* New York: Academic Press, 1973.

Hume, D. *Enquiry concerning human understanding.* London, 1748.

Hursh, S. R. The economics of daily consumption controlling food- and water-reinforced responding. *Journal of the Experimental Analysis of Behavior,* 1978, **29,** 475-491.

Kavanau, J. L. Behavior of captive white-footed mice. In E. P. Willems & H. L. Rausch (Eds.), *Naturalistic viewpoints in psychological research.* New York: Holt, 1969.

Killeen, P. Response rate as a factor in choice. *Psychonomic Science,* 1968, **12,** 34.

Krechevsky, I. Brain mechanisms and variability: I. Variability within a means-end-readiness. *Journal of Comparative Psychology,* 1937, **23,** 121-138.

Kulli, J. C., & Bogrow, P. A. A cumulative recorder for experiments on concurrent schedules. *Journal of the Experimental Analysis of Behavior,* 1971, **16,** 381-383.

Langer, E. J., & Rodin, J. The effects of choice and enhanced personal responsibility for the aged: A field experiment in an institutional setting. *Journal of Personality and Social Psychology,* 1976, **34,** 191-198.

Lefcourt, H. M. Internal versus external control of reinforcement: A review. *Psychological Bulletin,* 1966, **65,** 206-220.

Lefcourt, H. M. The function of the illusions of control and freedom. *American Psychologist,* 1973, **28,** 417-425.

Leigland, S. M. Deviations from matching as a measure of preference for alternatives in pigeons. *Journal of the Experimental Analysis of Behavior,* 1979, **32,** 1-13.

Locke, J. *Essay concerning human understanding*, London, 1690.

Lovitt, T. C., & Curtiss, K. Academic response rate as a function of teacher- and self-imposed contingencies. *Journal of Applied Behavior Analysis*, 1969, **2,** 49-53.

MacDonough, T. S., Adams, H. E., & Tesser, A. The effects of choice in systematic desensitization. *Psychological Record*, 1973, **23,** 397-404.

Menlove, R. L., Moffitt, M., & Shimp, C. P. Choice between concurrent schedules. *Journal of the Experimental Analysis of Behavior*, 1973, **19,** 331-344.

Mill, J. S. *On liberty*. London, 1859.

Monty, R. A., Rosenberger, M. A., & Perlmuter, L. C. Amount and locus of choice as sources of motivation in paired-associate learning. *Journal of Experimental Psychology*, 1973, **97,** 16-21.

Moore, J., & Fantino, E. Choice and response contingencies. *Journal of the Experimental Analysis of Behavior*, 1975, **23,** 339-347.

Munsinger, H., Kessen, W., & Kessen, M. L. Age and uncertainty: Developmental variation in preference for variability. *Journal of Experimental Child Psychology*, 1964, **1,** 1-15.

Navarick, D. J., & Fantino, E. Transitivity as a property of choice. *Journal of the Experimental Analysis of Behavior*, 1972, **18,** 389-401.

Neuringer, A. Delayed reinforcement *versus* reinforcement after a fixed interval. *Journal of the Experimental Analysis of Behavior*, 1969, **12,** 375-383.

Partridge, P. H. Freedom. In P. Edwards (Ed.), *The encyclopedia of philosophy* (Vol. 3). New York: Macmillan, 1967.

Perlmuter, L. C., Monty, R. A., & Kimble, G. A. Effect of choice on paired-associate learning. *Journal of Experimental Psychology*, 1971, **91,** 47-53.

Platt, J. R. Percentile reinforcement: Paradigms for experimental analysis of response shaping. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 7). New York: Academic Press, 1973.

Premack, D. Reinforcement theory. In D. Levine (Ed.), *Nebraska symposium on motivation*. Lincoln: University of Nebraska Press, 1965.

Premack, D. Catching up with common sense or two sides of a generalization: Reinforcement and punishment. In R. Glaser (Ed.), *The nature of reinforcement*. New York: Academic Press, 1971.

Rachlin, H., & Green, L. Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, 1972, **17,** 15-22.

Reese, E. P. *The analysis of human operant behavior*. Dubuque, Iowa: Brown, 1966.

Roeder, K. D. Neural factors and evitability in insect behavior. *Journal of Experimental Zoology*, 1975, **194,** 75-88.

Rozin, P., & Kalat, J. W. Specific hungers and poison avoidance as adaptive specializations in learning. *Psychological Review*, 1971, **78,** 459-486.

Schoenfeld, W. N., Harris, A. H., & Farmer, J. Conditioning response variability. *Psychological Reports*, 1966, **19,** 551-557.

Skinner, B. F. The phylogeny and ontogeny of behavior. *Science*, 1966, **153,** 1205-1213.

Skinner, B. F. *Beyond freedom and dignity*, New York: Knopf, 1971.

Skinner, B. F. The shaping of phylogenic behavior. *Journal of the Experimental Analysis of Behavior*, 1975, **24,** 117-120.

Skinner, B. F. Herrnstein and the evolution of behaviorism. *American Psychologist*, 1977, **32,** 1006-1012.

Steinman, W. M. Response rate and varied reinforcement: Reinforcers of similar strengths. *Psychonomic Science*, 1968, **10,** 35-36.

Taffel, S. J., & O'Leary, K. D. Reinforcing math with more math: Choosing special activities as a reward for academic performance. *Journal of Educational Psychology*, 1976, **68,** 579-587.

Toffler, A. *Future shock*. New York: Random House, 1970.

Voss, S. C., & Homzie, M. J. Choice as a value. *Psychological Reports,* 1970, **26,** 912-914.

Wicklund, R. A. *Freedom and reactance.* Potomac, Md.: Erlbaum, 1974.

Willis, R. D., Van Hartesveldt, C., Loken, K. K., & Hall, D. C. Motivation in concurrent variable-interval schedules with food and water reinforcers. *Journal of the Experimental Analysis of Behavior,* 1974, **22,** 323-331.

Zeigler, H. P., Green, H. L., & Siegel, J. Food and water intake and weight regulation in the pigeon. *Physiology and Behavior,* 1972, **8,** 127-134.

A SKETCH OF AN ECOLOGICAL METATHEORY FOR THEORIES OF LEARNING[1]

Timothy D. Johnston

NORTH CAROLINA DIVISION OF MENTAL HEALTH, RESEARCH SECTION
RALEIGH, NORTH CAROLINA

M. T. Turvey

UNIVERSITY OF CONNECTICUT, STORRS, CONNECTICUT AND
HASKINS LABORATORIES, NEW HAVEN, CONNECTICUT

---

[1] Preparation of this article was supported in part by Research Grant No. HD-00878 from the National Institute of Child Health and Human Development. The authors thank David B. Miller for his comments on a portion of the manuscript.

ISBN 0-12-543314-X

I. Introduction

For about the first half of this century, the psychology of learning was unified by a set of metatheoretical concepts and beliefs that may loosely be termed the general process view of learning (Seligman, 1970). Although the major general process theorists (Pavlov, Thorndike, Watson, Guthrie, Tolman, Hull, Spence, and Skinner) differed sharply on a number of theoretical issues, they all shared a set of common assumptions about learning that allowed those issues to be clearly defined and that enabled workers in the field to agree on the nature of important questions to be asked about learning. The issues that were debated included those of S-S versus S-R learning, reinforcement versus contiguity, the nature and role of drive states, and the importance of cognitive processes in learning. Underlying these theoretical arguments was a common set of metatheoretical beliefs that, within the general process tradition, were not called in question. These included the belief that there are general principles of learning that apply to all learning situations, that the same learning processes are involved in all animals, and that learning is to be equated with the formation of associations of some kind.

The last 25 years have seen a gradual retreat from the general process view of learning as its underlying assumptions began to be questioned. The view that learning is explicable by a single set of general principles has given way to a belief in a multiplicity of principles, a development that was anticipated by Tolman (1949). Thus we have seen a proliferation of "minitheories," each dealing with a restricted range of learning phenomena such as classical conditioning (Rescorla, 1972; Rescorla & Wagner, 1972), expectancy (Kamin, 1968, 1969), discrimination learning (Mackintosh & Sutherland, 1971), and avoidance learning (Bolles, 1970, 1971, 1972). The phylogenetic generality of learning processes has been questioned by proponents of the "biological boundaries" approach to learning (Bolles, 1970; Kalat, 1977; Rozin & Kalat, 1971; Seligman, 1970; Shettleworth, 1972). In the literature on human learning in particular, associationism is on the wane and cognitive theories of learning are preeminent (Haugeland, 1978).

The current state of the psychology of learning, in short, is one of great conceptual diversity in which little attention is being paid to the prospects for a unified approach to the study of learning. It may perhaps be argued that such diversity is in fact just what is required to free the study of learning from the remnants of the general process view and to provide a broader data base on which to build new theories. An alternative argument, however, is that without a competing metatheory to set against the general process view, theoretical unification of the field will be greatly retarded, because there will be no basis for agreement on the theoretical issues to be resolved and on the important questions that must be asked (see Kuhn, 1962, for arguments and examples supporting this philosophical position). Our aim in this article is to sketch the outline of such a competing metatheory. To do this, we have chosen to stand outside the mainstream of much current psychology of learning and to consider learning within the conceptual framework of evolutionary and ecological biology. The biological boundaries approach to learning has set an important precedent for taking ecological considerations seriously in the study of learning, but we shall attempt to offer a more radical alternative to general process theory than is provided by the latter approach (see Johnston, submitted).

The distinction between *metatheoretical* analysis, which primarily concerns us in this article, and *theoretical* analysis, with which we shall be more peripherally concerned, is perhaps worth making explicit. Briefly, the distinction is this: Theoretical analysis provides answers to questions that are posed on the basis of metatheoretical analysis. Metatheory is concerned with justifying the asking of certain kinds of questions in a particular area of inquiry, in our case, the area of learning, and putative answers to those questions are presented in the form of theories, hypotheses, and models. One may arbitrate among competing theories by pointing to data that conflict with some but not with others, but competing metatheories cannot be evaluated in this way. Asking certain kinds of questions about learning may be unprofitable but the questions asked are not "wrong" in any factual sense. A metatheory can thus be properly evaluated only after some of the research that it mandates has been carried out, so that its profitability may reasonably be assessed. In the interim, a preliminary judgment may be made on the basis of criteria such as how well the metatheory articulates with those of other, related fields of inquiry, how wide a range of phenomena is opened up for analysis under the metatheory, and whether the questions asked seem to be prima facie amenable to experimental study. These criteria are in addition to those of a logical kind that apply to any form of intellectual inquiry, such as logical consistency and coherence, a minimum number of unproven assumptions (which should be unproblematical), precise definition of central terms, and so forth.

The distinction between metatheory and theory has much in common with Kuhn's (1962) distinction between a paradigm and the normal science that it sanctions, and the concept of metatheory is very similar to Lakatos' (1970) "scientific research programme." This is not, however, the place for a detailed comparison of these various concepts. The aim of this brief philosophical digression has been to characterize the nature of our endeavor and to suggest partial criteria by which it might reasonably be judged.

II. The Nature of Ecological Inquiry

Ecology, as a branch of science in its own right, studies the relationships between living organisms and the world in which they live. Thus defined, ecology encompasses almost all of the disciplines that include an organism as part of their concerns, but tradition and necessity have combined to give the field a somewhat more restricted scope (Elton, 1927; Emlen, 1973; Odum, 1953). The *ecological approach,* however, is one that may usefully be applied to any of a wide range of problems that lie outside the scope of *ecology* as traditionally defined. Taking our cue from the above definition, we may say that an ecological approach is one that studies some aspect of an animal[2] in relation to the environment that it inhabits, or, in complementary fashion, that studies some aspect of an environment in relation to the animal that lives in it.

A. ANIMAL AND ENVIRONMENT

The nature of ecological inquiry depends, evidentally, on how one construes the relationship between an animal and its environment; and the nature of this relationship will depend, again, on the interpretation placed on the two terms, ''animal'' and ''environment,'' that enter into it.

1. The Animal as Actor

The conventional and intuitive interpretation of the term ''animal'' is as a morphologically defined entity, bounded by an epidermis that sets it off from the rest of the world, conventionally its ''environment.'' As long as our interest in the animal is appropriately pursued in morphological terms, such an interpretation may be appropriate. The study of learning, however, is not concerned with questions of morphology, but rather with questions of behavior and of change in behavior, and so we may question the suitability of a morphological interpretation of the term ''animal'' for the study of learning.

---

[2] Throughout this article we will be using the terms ''animal'' and ''actor'' in the sense of ''species-typical individual.'' That is, we will be thinking of the animal or actor as possessing certain definite (though perhaps unspecified) characteristics that make it typical of a species or population. The phrase ''individual animal (actor)'' will be used when we wish to denote a particular individual, which may or may not be species-typical.

In morphological terms, we might observe that an animal may be described as possessing certain structures, such as limbs, organs, tissues, and cells. In behavioral terms, then, let us observe that an animal may be described as effecting certain actions, such as feeding, walking, rasing its head, and moving its eyes. Thus we may say that from a behavioral point of view of animal is an *actor,* defined in terms of a set of *effectivities* (Turvey & Shaw, 1979), that is, of actions that effect consequences for the actor.

In choosing the term "actor" in preference to "animal" our aim is draw attention to the fact that more than one kind of description may be given of an organism. The description that we choose must be appropriate for the kind of analysis we wish to pursue and, having chosen a particular style of description, it is important not to confuse its terms with those of other styles of description. Such confusion leads to what philosophers call "category errors," in which properties appropriate to the elements of one style of description are inappropriately applied to those of another style. The usual result of such errors is that questions are asked that are unhelpful and misleading: "What color is the Law of Effect?" is a blatant example of a category error. A more famous example is provided by Gilbert Ryle's (1949) foreign visitor to Oxford who was shown the various colleges and who then asked to be taken to the University; "colleges" and "the University" are elements of different styles of description of Oxford and only colleges can be assigned the property of physical location within the town. More subtle, and hence dangerous examples from the study of behavior are discussed by Fodor (1968) and by Purton (1978).

Behavior and morphology are alternative styles of description of an organism. By selecting a term such as "actor," which is both descriptive and relatively unfamiliar in this context, to apply to the results of a behavioral description, we hope to keep the two styles clearly distinct and so avoid unintentionally committing category errors.

2. *The Environment as Econiche*

Under the conventional definition of "animal" as a morphological entity, "environment" is construed as a physical entity, namely, that part of the world outside the animal's skin. Defining the animal as an actor, in terms of a set of effectivities, raises difficulties for this definition, however, for an effectivity is a description of an actor *relative* to some environment. The effectivity of flight, for example, can be realized by a particular actor only under certain circumstances (which will vary between actors) and to say that an actor possesses the effectivity of flight is necessarily to imply an environment in which those conditions are realized.

A description of an animal as an actor possessing a particular set of effectivities, or an *effectivity structure* (Turvey & Shaw, 1979), identifies, in short, an *econiche,* which is an environment possessing the necessary support for those effectivities. A description of the ecological support for an effectivity defines an *affordance* (Gibson, 1977; Turvey & Shaw, 1979), which is a specific combination of physical properties of an environment taken with reference to a particular effectivity. A description of the physical properties of an environment, taken with reference to the effectivity structure of an actor, defines the *affordance structure* of an econiche for that actor. It is important to appreciate the significance of the phrase ''for that actor'' since a structure that affords climbing, say, by an actor of one kind (such as a tree frog) may not afford climbing by some other actor (such as an alligator).

The term ''econiche'' is derived from the ecological concept of the niche, originated by Grinnell (1924) and greatly elaborated by Hutchinson (1957, 1967; see Vandemeer, 1972, and Whittaker, Levin, & Root, 1973, for more recent discussions). As proposed by Hutchinson (1967), the niche is a volume of *n*-dimensional hyperspace, each of the dimensions corresponding to some physical factor of ecological relevance to the organism; points within the niche space define conditions under which the organism can survive. The concept of econiche, as we shall employ it, is limited to factors of behavioral relevance to the animal and, in particular, it is an animal-relevant description of the environment.

The concept of affordance is discussed at greater length in Section IV,B. For the moment, we wish to emphasize some of the implications of the preceding discussion. It will be apparent that effectivities and affordances are complementary descriptions of actors and econiches, respectively, taken with respect to each other. Just as describing an actor in terms of its effectivity structure implies an econiche with a particular affordance structure for it to inhabit, so describing an econiche with a particular affordance structure implies an actor with the requisite effectivity structure to inhabit it. This statement must not be read as a claim that the existence of the physical world is dependent on the existence of animals; as previously remarked, the term ''econiche'' refers to a *description* of the physical world with respect to some actor. ''Actor'' and ''econiche,'' as we shall use these terms, are coimplicative and cannot be defined independently of one another. Figure 1 illustrates the relationships among the various terms defined in this section.

This discussion by no means exhausts the problems inherent in the concept of the econiche; indeed the development of an adequate theory of

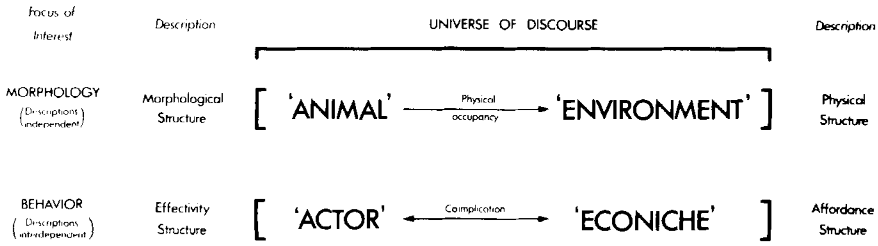


Fig. 1. Two styles of description of the animal and the environment, one from a morphological point of view, in which the descriptions given are independent, and one from a behavioral point of view, in which they are interdependent or coimplicative.

environments taken in relation to their inhabiting animals remains an important desideratum for ecological theory in general (Mason & Langenheim, 1957). For the present, we will content ourselves with supplementing the above account with two further considerations. First, since the term "econiche" refers to a description rather than to a locality, any specific locality will not necessarily provide all of the affordances required by a particular actor. We may speak, therefore, of a locality as providing an *econiche for S,* where *S* is some subset of the effectivities of an actor. There must be a set of localities, however, such that together they constitute an econiche for the actor and such that they are *connected* for the actor. Two localities are connected for an actor if it possesses an effectivity permitting it to move from one locality to the other.

The second consideration involves the problem of ontogeny, which is obviously of prime importance in any discussion of learning. As the actor develops, its effectivity structure changes. It acquires capabilities that it previously did not possess and it loses others. Concurrently, the affordance structure of its environment changes. This does not mean, of course, that the physical characteristics of the localities it occupies change (although they may do), but that those characteristics change in relation to the changing effectivity structure of the actor. Mason and Langenheim (1957) express this important point as follows: "The life-span of the organism is the duration time of its environment [econiche] [and] . . . the environmental relation [affordance structure] is ordered by the ontogeny of the organism" (pp. 331-332).

B. THE ECOSYSTEM—UNIT OF ECOLOGICAL INQUIRY

The close and complementary relationship between actor and econiche that is established by our analysis means that the focus of inquiry into learning must broaden to include more than just the animal. The ecological approach to learning is concerned with mutually defined actors and econiches and so the minimal unit of our inquiry must be the whole actor-econiche system, which we will term the *ecosystem*. This term was proposed by A. G. Tansley in 1935 and has become a prominent component of ecological theory, especially in the work of Odum (1953, and subsequent editions). Our use of the term differs somewhat from that of Tansley and of most subsequent writers but it preserves the essential ingredient of their thinking: that of an integral system of animal and environment.

Odum (1959, p. 10) defines the ecosystem as follows:

> Any area of nature that includes living organisms and nonliving substances interacting to produce an exchange of materials between the living and nonliving parts is an ecological system or ecosystem.

As examples of ecosystems, Odum (1959, p. 11) lists "a pond, a lake, a tract of forest or even a small aquarium." In ecological theory, the various organisms that inhabit a locality such as a lake comprise a community and the term "ecosystem" as used by most ecologists refers to the community and its environment. We will use the term however to refer to a single (kind of) actor and its econiche, as we have previously defined those terms.

The utility of allowing the concept of ecosystem to apply at various scales of biological organization (individual, group, population, species, or community) was pointed out by Evans (1956), who also stressed that the ecosystem, rather than any of its components, is the minimal unit of ecological inquiry. Where necessary, we shall adopt Evans' suggestion that "the particular level on which the ecosystem is being studied can be specified with a qualifying adjective—for example, community ecosystem, population ecosystem, and so forth" (p. 1128). Since our main concern is with individual kinds of actors, the unadorned term "ecosystem" will be used for this scale of analysis.

The focus of our ecological approach is somewhat more restricted than that of the traditional ecologists, for our concern is with the ecological support for behavior and, crudely, with the ways in which an actor learns to exploit that support in the course of its development. The more traditional issues of the flow of matter and energy through the ecosystem will not concern us.

C. A STYLE OF INQUIRY FOR THE ECOLOGICAL STUDY OF LEARNING

> What and where is behavior? The location of behavior is literally in naturally evolving life on earth. It is literally in organism-environment. These sentences are not verbal

> generalities about some generality of behavior. They are intended as literal report upon the specific instance of specific behavior. (Bentley, 1941, p. 485)

In putting forth the ecosystem as the minimal unit for the ecological analysis of learning, we make the same claim for learning as Bentley makes for behavior in the above quotation: Learning goes on in ecosystems, not in animals. The ecological approach adopts a style of inquiry that Dewey and Bentley (1949) call ''transactional,'' a style that, ''assumes no pre-knowledge of either organism or environment as adequate . . . but requires their primary acceptance in common system'' (p. 123). This style of inquiry Dewey and Bentley (1949) contrast with ''interactional'' inquiry, which ''assumes the organism and its environment objects to be present as substantially separate existences or forms of existence, prior to their entry into joint investigation'' (p. 123).

Traditional approaches to the study of learning have adopted an interactional style of inquiry. The animal is described as the possessor of a set of responses and the environment as the emitter of a set of stimuli, each of these sets being described independently of the other, and then means are sought of mediating between one set and the other by recourse to constructs such as reinforcement, contiguity, expectancy, and so forth. By contrast, we seek a style of inquiry that respects the integrity of the ecosystem and that exploits the coimplicative relationship between actor and econiche in an attempt to dispense with the need for mediation between the two (see Shaw & Turvey, in press).

In pursuit of this end, we require a system of concepts that will permit us to treat the integrity of the ecosystem as a primary datum, not as derived from an interaction between animal and environment. This is the task to which we turn in the following section.

III. The Nature of Biological Adaptation

Having described the nature of ecological inquiry, we now focus on the nature of the relationship between actor and econiche that is the crucial element in any such inquiry. In the course of this discussion we shall consider a number of important concepts that provide the conceptual tools required for the ecological analysis of learning.

A. THE CONCEPT OF ADAPTATION

Central though it is to modern biological theory, few concepts have generated more confusion or proven more resistant to analysis than that of adaptation (Ghiselin, 1966; Medawar, 1951; Williams, 1966). We will not attempt to provide a full-scale explication of the concept here, since to do so would carry us well beyond the scope of this article. Rather, we shall provide a more limited discussion, aimed at resolving certain issues that are particularly relevant to arguments that we wish to make later. The reader interested in a more complete discussion may wish to consult some of the very large literature on the topic, of which the following may be cited as especially valuable: Bock and von Wahlert, 1965; Dobzhansky, 1942, 1956, 1968; Ghiselin, 1966; Lewontin, 1978; Medawar, 1951; Slobodkin, 1968; Slobodkin and Rapoport, 1974; Sommerhoff, 1950, 1969; Williams, 1966.

In the first place, we must distinguish between *adaptation* and *fitness,* two closely cognate terms that are frequently, though incorrectly, used synonymously (for example by Lewontin, 1956, and by Stern, 1970; cf. Dobzhansky, 1956, 1968; Ghiselin, 1974). Fitness is the more easily defined term, being the relative reproductive contribution that an individual makes to the next generation, in comparison with that of other individuals in the population, under a defined set of environmental conditions (Lewontin, 1974; Mettler & Gregg, 1969). It is differences in reproductive fitness that give rise to natural selection (Darwin, 1859), since in a stable population, those individuals that contribute most offspring to subsequent generations will increase their genetic representation at the expense of those that contribute least.[3] There is a close and nonarbitrary relationship between adaptation and fitness, and we discuss this relationship in Section IV,A.

By contrast with the concept of fitness, adaptation is to be assessed more in terms of individual *survival* than in terms of individual *reproduction.* We shall first develop the concept of adaptation with regard to the animal and its environment, after which it will be seen that a transition to the concepts of actor and econiche may be accomplished quite naturally. If an animal *A* can survive in an environment *E,* then we may say that *A* is adapted for survival in *E*. This formulation is acceptable as far as it goes but it does not go very far; it leaves ''survival'' as a primitive, unanalyzed term, but it is the means whereby survival is ensured that must concern us. If we are to make sense of learning in the context of

---

[3] This formulation is somewhat inaccurate, for it neglects the concept of inclusive fitness (Hamilton, 1964), which has come to play an important role in modern evolutionary theory. If we ride roughshod over such subtleties (as we do here and in other parts of this article), it is not because of a lack of appreciation of their theoretical importance, but rather out of sympathy for our psychological readership, whose interest is primarily in learning rather than in evolution. We have not, of course, adopted this cavalier attitude when to do so would introduce distortions into the substance of our arguments.

biological adaptation then we require a formulation that recognizes the complexity of animal–environment relationships underlying survival and that will allow us to elucidate the role of learning in maintaining those relationships. Let us see how this might be accomplished, adopting a strategy based on that originated by Sommerhoff (1950).

One way of expressing *A*'s adaptedness for survival in *E* is to say that *A* can attain the goal of survival in *E*. With this formulation, the way is open for us to unpack the concept of adaptation: Let us define survival as an ultimate goal of *A*, in the sense that all of *A*'s activities are adaptively significant only to the extent that they contribute to survival. There are then a number of subgoals, varying in detail between animals of different kinds, that must be attained if survival is to be assured. Thus *A* must be able to obtain food, avoid predators, move through space, orient, acquire a mate, care for its young, and so forth. Each of these requirements defines a goal that contributes to eventual survival and reproduction. If *A* is able to attain a goal *G* under some set of conditions *E*, then we may say that *A* is adapted for *G* in *E*.

Defining adaptation in terms of the attainment of goals does not represent a retreat into either vitalism or teleology. "*A* can attain goal *G* under conditions *E*" is a descriptive, not an explanatory statement and it is hence not teleological. We must still account for the fact of goal attainment and such an account need not (and should not) appeal to an *élan vital* or other inexplicable entity. Rather, an explanation must be given in terms of publicly observable characteristics of the goal-directed system, in terms of its "objective system properties" (Sommerhoff, 1950, 1969). Equally important, the concept of goal-directedness does not require us to assume that *A* is conscious of its goal, that it is acting purposefully, or indeed that it has any internal representation at all of the goal state. Any of these conditions *may* hold, but they need not. What is important for the animal is that it be able to attain the goal—articulating or representing the goal is a problem for the theorist, along with the deeper problem of explaining the process of goal attainment. We shall have more to say of these problems, especially the latter, in the following section.

In defining *A*'s adaptedness in terms of both *G* and *E* we have given explicit recognition to an important aspect of the concept of adaptation, namely, that it is a *relational* concept (Bock & von Wahlert, 1965; Slobodkin & Rapoport, 1974; Sommerhoff, 1950). An animal does not possess adaptation in the sense in which it possesses mass, length, or limbs. Rather, it may stand in an adaptive relationship to some environment and the environment must be defined in order for us to discuss the animal's adaptation. It makes no sense to say that *A* is adapted for some *G* without specifying the environment *E* for which such adaptation holds.

Adaptation (more precisely, adaptation for some *G*) is a property of *ecosystems,* of animal-environment systems, not of animals alone.

The definition of adaptation in terms of goal attainment makes it undesirable to speak, as did Henderson (1913), of an environment as also being adapted to the animal that inhabits it (although we are naturally sympathetic to the ecological tenor of Henderson's thesis). Since we now wish to incorporate the mutually defined concepts of actor and econiche into our analysis, however, we must have some way of referring to the complement of the adaptive relationship of the animal to its environment. We shall speak, therefore, of the *appropriateness* of an environment for an animal that is adapted to it (by implication, in relation to some goal *G*). It is just as legitimate, although it may be less familiar, to inquire into the source of an environment's appropriateness as it is to inquire into the nature of an animal's adaptation. In the latter case we seek to understand the biological characteristics that enable *A* to attain *G* in *E;* in the former, we seek to explain the ecological support that *E* provides for the attainment of *G* by *A*. Note that appropriateness is a relational concept in the same way as adaptation: An environment is not appropriate per se, but only in relation to the animal that is adapted to it. It will now be apparent that the biological characteristics of interest are the effectivities that permit us to describe *A* as an actor and that the ecological support is provided by the affordances that allow us to describe *E* as an econiche. Figure 2 illustrates the complementary relationship between an adapted animal (the actor) and its appropriate environment (the econiche).

Several authors (e.g., Medawar, 1951; Stern, 1970) have pointed out that the term "adaptation" may be used in a number of different senses. To avoid confusion, we will define the three most important senses as follows:

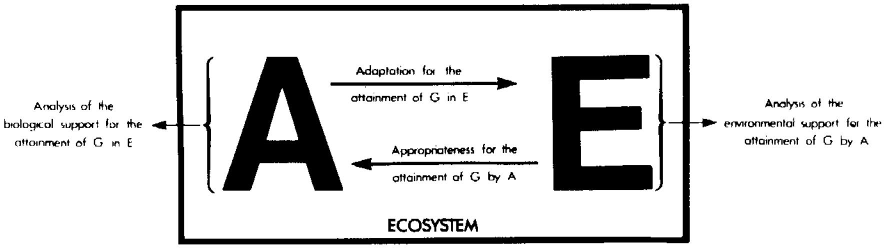


Fig. 2. The actor *(A)* and its econiche *(E)* together constitute an integrated system, the ecosystem. The relationship of *A* to *E* is one of adaptation for the attainment of goals *(G);* the relationship of *E* to *A* is one of appropriateness for the attainment of those goals. Both the actor's adaptation and the econiche's appropriateness may be analyzed in terms of the support they provide for goal attainment.

1. As a *relation*. This is the sense in which the term has been used in the preceding discussion;
2. As a *characteristic*. An adaptation, or adaptive feature, is some characteristic of A that enables it to survive (and reproduce) in E. An adaptation for some goal G (e.g., a "feeding adaptation") is one that enables A to attain G in E. Determining whether a particular characteristic of an organism is in fact an adaptation, and if so in what way, is an extremely difficult problem for which there is no general solution (Bock & von Wahlert, 1965; Hinde, 1975; Lewontin, 1979);
3. As a *process*. An adaptive process is one that gives rise to an adaptive relationship between A and E in regard to some G. Adaptations [in sense 2 above] are a product of (a process of) adaptation.

The aim of our analysis is to provide an understanding of learning in relation to adaptation construed in all three senses. First, we want to understand learning as a *process* of adaptation that is manifest over relatively short periods of time, within the lifespan of an individual actor. Second, we wish to understand learning as itself a *product* of adaptation, in this case of the process of evolutionary adaptation, acting over much longer periods of time. Finally, our analysis of both of these aspects of the problem will be guided by the nature, already outlined, of the adaptive *relationship* between the actor and its econiche, for it is the maintenance of this relationship that is the *raison d'être* for learning as both product and process.

B. A FORMAL MODEL OF GOAL ATTAINMENT

Having provided, in the preceding section, an account of adaptation as a relationship, we now turn to an account of adaptation as a process. Here we shall be concerned not with particular kinds of adaptation, such as learning, nor with the problem of how adaptation occurs, in the sense of providing hypothetical mechanisms. Our analysis remains in the domain of metatheory, in that we are concerned with the question of *what* adaptation is as a process, seeking to answer that question in precise and formal terms. Once the formal model has been expounded, we shall see that it has important implications for the subsequent analysis of learning in relation to adaptation.

The model we shall present is a simplified version of one that has been worked out in detail by Sommerhoff (1950, 1969). Our account of it will necessarily be brief; we will emphasize those aspects of the model that are of greatest relevance to our present concerns and, in particular, we will omit much of Sommerhoff's mathematical development. We will develop the model with respect to an example of goal-directedness from the field of insect behavior, namely, prey capture by the praying mantis. The mantis catches small insects by means of its long, clawed forelimbs, a process that has been described in some detail by Maldonado, Levin, and Barros Pita (1967), Mittelstaedt (1957), and others. We will deal here with only a few selected aspects of this behavior. Figure 3 shows the position of the mantis' forelimbs at two instants: just before the strike (time = t_0) and just after the strike (time = t_k). Consider two variables describing the relationship between the mantis and its prey: the bearing of the prey (ϕ) and the bearing of the claw tip (β), both taken with regard to an arbitrary line drawn through a fixed point on the mantis' body (Fig. 3). These variables take the values ϕ_0 and β_0, respectively, at t_0 and ϕ_k and β_k at t_k.

The mantis' claw tip must be aligned with the prey in order for capture to be successful and so we may say that the goal of prey capture will be attained only if the following condition holds just after the strike (Fig. 3):

$$\phi_k - \beta_k = 0 \tag{1}$$

Equation (1) defines the *focal condition* of adaptation; it specifies the condition that must be satisfied if the goal is to be attained. In this example there will be many values of ϕ_0 (the initial bearing of the prey) for each of which the mantis can produce a specific, adapted value of β_k at t_k satisfying the focal condition. This range of values defines a set S_0, which may be either an interval on a continuum or a set of discrete values. Sommerhoff (1950) claims, with some justification, that S_0 must have at least two members in any instance of adaptation. However, we will consider some examples below (see Section V,A) in which there is but one member of S_0 and yet which are incontestably examples of adaptation.

In the case of the mantis, S_0 includes a range of values of ϕ_0 and we may say that the mantis possesses a "strike aiming system" such that on detecting the value of ϕ_0 at t_0 it produces a corresponding, or adapted value of β_k at t_k such that $\phi_k - \beta_k = 0$. It is the correspondence of ϕ_k and β_k that defines the adaptiveness of the aiming system and this correspondence is effected by virtue of a specific sensitivity of the system in regard to the value of ϕ_0. Following Sommerhoff (1950), we will refer to ϕ_0 as the *coenetic variable* of adaptation.

In the example we have been discussing, the relationship between ϕ_k and β_k that defines the focal condition is one of simple equality. Suppose, however, that ϕ is the size of the prey and β is the angle of the claw. Then there will be some more complex relationship between ϕ_k and β_k that must be satisfied if the prey is to be grasped securely and the goal of prey

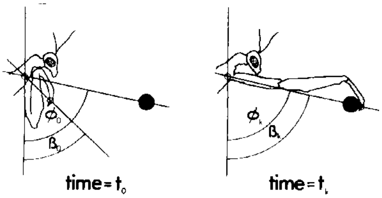


Fig. 3. Two instants in the predatory strike of a praying mantis, just before the strike (time = t_0) and just after the strike (time = t_k). The bearing of the prey (ϕ) and of the claw tip (β) are equalized as a result of adaptation over the interval $t_k - t_0$.

capture attained (see Holling, 1964, for details). We may describe this relationship by some function F and rewrite the focal condition in Eq. (1) more generally as:

$$F(\phi_k, \beta_k) = 0 \tag{2}$$

So far, we have considered only one variable describing the environment (ϕ) and one describing the animal (β), but in many instances of goal-directedness, several variables describing the animal ($\beta_1, \beta_2, \ldots, \beta_n$) must be adapted to several variables describing the environment ($\phi_1, \phi_2, \ldots, \phi_m$). For example, not only must the bearing of the prey and of the claw tip be equalized by the striking mantis, but the claw tip must also be positioned at an appropriate distance, just behind the prey (see Fig. 3). Thus Eq. (2) may be written still more generally as:

$$F(\phi_{1k}, \phi_{2k}, \ldots, \phi_{mk}, \beta_{1k}, \beta_{2k}, \ldots, \beta_{nk}) = 0 \tag{3}$$

To simplify discussion, we will employ the notation of Eq. (2) and allow ϕ and β to stand for any of several variables that we may wish to consider. Since we are concerned only with the formal and not the quantitative nature of the relationships between the animal and the environment, this simplification will not limit the generality of our arguments.

Two assumptions have been made in the preceding discussion: first, that the coenetic variable (ϕ_0) and the variable to which β_k is adapted (ϕ_k) are the same; and second, that the environment does not change between t_0 and t_k (i.e., $\phi_0 = \phi_k$). We may relax both assumptions, increasing the power and generality of the model. In many instances of biological adaptation, as we shall see, the coenetic variable may be different from ϕ, the variable to which adaptation is effected at t_k. For example, diapause in insects is a physiological adaptation to cold weather, but it is initiated as a response to short day length, not to low temperature (Beck, 1968). Here the coenetic variable is day length and the adaptive relationship holds between temperature (high or low) and the insect's physiological state (normal or diapause). In the sequel, we shall let y stand for the coenetic variable, taking the value y_0 at t_0. The nature of the insect's environment is such that between t_0 and t_k short day length (y_0) maps on to low temperatures (ϕ_k) and the constitution of the insect is such that, in the same interval, short day length maps on to diapause (β_k).

Formally, we may define two functions, one *(P)* defined on the environment, the other *(B)* defined on the animal, such that:

$$P(y_0) = \phi_k \tag{4}$$

and

$$B(y_0) = \beta_k \tag{5}$$

Notice that since the value of ϕ at t_0 (ϕ_0) does not appear in Eqs. (2)–(5), we may relax the assumption of an unchanging environment. This is replaced by the assumption that there is a coenetic variable y (possibly though not necessarily equal to ϕ) satisfying Eqs. (4) and (5). Notice also that the value of β_0 does not appear in Eqs. (2)–(5), implying that the initial state of the adapting organism does not affect the process of adaptation. This may be a limitation of the model, especially in regard to adaptive processes involving evolution and individual development. In these cases, the initial state of the system (at t_0) may be of crucial importance in determining the range of subsequent states that it can attain (at t_k)

The characteristics of a goal-directed, or adaptive system may then be summarized as follows (see Sommerhoff, 1969, pp. 174–175):

1. At some time t_k it is a necessary condition for the subsequent occurrence of a goal event G that the two sets of variables $\phi_1, \phi_2, \ldots, \phi_m$ and $\beta_1, \beta_2, \ldots, \beta_n$ should satisfy the focal condition in Eq. (3).
2. There is a coenetic variable y and two sets of functions $P_1, P_2 \ldots, P_m$ and $B_1, B_2, \ldots, B_n$ such that:

$$P_i(y_0) = \phi_{ik} \qquad i = 1, 2, \ldots, m \tag{6}$$

$$B_i(y_0) = \beta_{ik} \qquad i = 1, 2, \ldots, n \tag{7}$$

where y_0 is the value of y at t_0.

3. There is a set S_0 of values of y_0, often containing at least two members, but sometimes only one, for which the above conditions hold.

Figure 4 presents a diagrammatic representation of the various elements in Sommerhoff's model, in the form we shall employ in subsequent discussion. This figure may be referred to when the model is employed in the analysis of particular examples of adaptation in Sections IV and V.

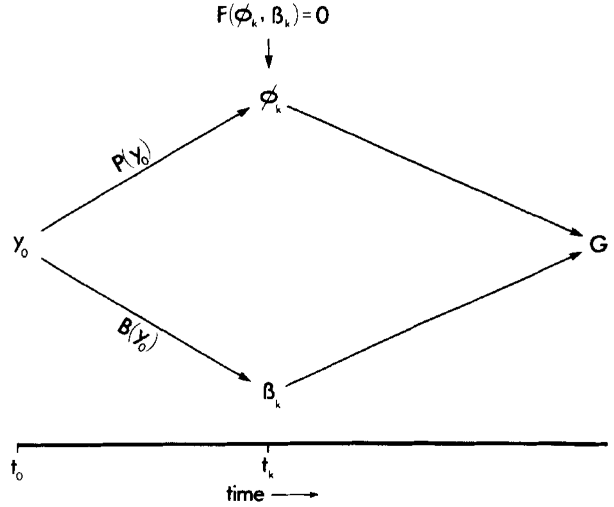


Fig. 4. Sommerhoff's formal model of goal attainment. See text for discussion.

C. THE BACK-REFERENCE PERIOD

We now single out one element of the model for special consideration, since it will play an important part in subsequent discussion. This is the interval $t_k - t_0$, called by Sommerhoff (1969) the *back-reference period*. The back-reference period is the operation lag of the adaptive mechanism that implements the response β_k. During this period, two events occur: the function B maps y_0 on to β_k and the function P maps y_0 on to ϕ_k. In order for β_k to be an adaptive response to ϕ_k at t_k, the function P must be determinate and single valued; that is P must always map a given value of y_0 on to the same value of ϕ_k. If this were not the case, B would often map y_0 on to a maladaptive value of β_k, since the value of ϕ_k would be inappropriate. Another way of saying this is that *given an adaptive mechanism with a back-reference period* $t_k - t_0$, *y and* ϕ *must be invariantly related over at least the span* $t_k - t_0$. Alternatively and equivalently, *given an environment such that y and* ϕ *are invariantly related over a span of no more than* $t_k - t_0$, *adaptation to* ϕ *can be effected only by an adaptive mechanism with a back-reference period of* $t_k - t_0$ *or less.*

This is a formal expression of a boundary condition on the relationship between a variable environment and an adapting animal: Different rates of variation, measured by the invariance between y and ϕ, require adaptive responses with different back-reference periods. Table I illustrates this point with reference to four kinds of variability and the adaptive response

TABLE I

THE CORRESPONDENCE BETWEEN ADAPTIVE RESPONSES AND THE TIME-SCALE OF ENVIRONMENTAL VARIABILITY

Kind of variability	Back-reference period	Adaptive response
Obstacles in the path of locomotion	Fraction of a second	Adjust direction of movement
Appearance of predator	Few seconds	Effect avoidance or defensive behavior
Location of food, water, shelter, etc.	Several days to a few years	Learn appropriate routes of travel
Availability and nature of substrates for locomotion	Many years	Evolution of locomotor system by natural selection

that is appropriate to each. It is the time-scale of the change and the back-reference period of the response that determine whether the response can be adaptive with regard to the change.

D. UNIFICATION IN A THEORY OF ADAPTIVE RESPONSE

It has been pointed out before, for example by Plotkin and Odling-Smee (1979), Slobodkin (1968), and Slobodkin and Rapoport (1974), that animals must be able to adapt to environmental change occurring on different time-scales and that in order to do this, different adaptive mechanisms are required. Perceptuomotor coordination, learning, and evolution by natural selection are three of the most prominent such mechanisms, effective in regard to short-term, medium-term, and long-term environmental change, respectively. A unified theory of adaptive response would provide an account of adaptation general enough to cover all of these (and other) cases, while at the same time providing an explicit statement of the differences between different styles of adaptive change. Sommerhoff's model provides just the framework for such an account and in Section IV,C we shall employ it in our discussion of learning as a mode of ecological adaptation.

The generality of Sommerhoff's model lies in its explicit formalization of several crucial elements in any form of adaptive response: a goal event (G); a focal condition $[F(\phi_K, \beta_K) = 0]$; a coenetic variable (y); one (or more) environmental variables (ϕ); and one (or more) response variables (β). In addition, it provides grounds for explicit distinctions between different forms of adaptive response. We have stressed the back-reference period $(t_K - t_0)$, but it should also be realized that the function mapping y_0 on to β_K [i.e., $B(y_0) = \beta_K$], which corresponds to the *mechanism* of adaptive response, is a further important basis for distinction. The model leaves many of these elements unanalyzed for, as Sommerhoff (1969) remarks, he is concerned primarily "with the general aspect of *what* a [goal-directed] system does and not *how* it does it" (p. 152). That is also our concern in this metatheoretical article. As remarked in Section I, we wish to motivate the asking of certain kinds of questions about learning by presenting the case for a particular, ecological account of *what* learning is. The answers to some of those questions will constitute putative accounts of *how* learning occurs.

The generality that is provided by this model is bought at the expense of certain restrictions on the account that we can give of particular kinds of adaptive response. The most important of these restrictions is that by accepting the generality of the model, we deny ourselves the option of setting up different forms of adaptive response, such as perception, learning, and evolution, as distinct categories, as phenomena *sui generis*, each to be accounted for on different principles. Instead our strategy must be to identify the various elements of the model in each instance of adaptation and to use the model to explicate the relationship among those elements. This does not mean that we see no important differences among the various forms of response, or that we see no opportunity for the separate development of theories of perception, of learning, and of evolution. We do not deny that important differences exist, particularly in regard to the mechanisms of response, but we claim that much is to be gained by seeking to construct theories of different forms of response that are closely compatible with each other.

The question may be raised whether the unified account at which we aim is an accurate reflection of reality. We believe that it is. If we take any temporal scale of description of the environment, from milliseconds to millenia, we observe that at each scale some aspects remain invariant while others vary. Insofar as these aspects of the environment describe an econiche for some actor, adaptation to both variant and invariant features at all relevant scales of description must be achieved. Such adaptation cannot be achieved by incompatible processes operating at odds with one another; adaptation must be an integrated response by the entire biological system and it must be effected concurrently at all levels of biological organization. This integration of response is perhaps the most remarkable feature of biological systems and it is often obscured by the necessity of isolating particular responses for experimental examination. One of the major goals of biological and psychological theory must be to account not just for particular forms of response, isolated in the laboratory, but for the overall *unified* adaptation of the organism to its environment. Such a goal is more likely to be realized if a unified approach to the problems of adaptation is adopted at the outset than if it is attempted only when several disparate theories have already established themselves.

Adopting a unified approach to the various forms of adaptation is consonant with an epistemological position that we might call "pragmatic realism" (see Shaw, Turvey, & Mace, in press). Adaptation, of whatever kind, is effective only if it works in some real environment. Consider the case of evolutionary adaptation: Evolutionary change works as an adaptive response only if it is a change in regard to some real feature of the environment actually inhabited by the animal. It would be absurd to maintain that natural selection effects adaptation to some *representation* of the environment and that this representation relates only equivocally to the real environment, yet this is the position adopted by the majority of current theories of short-term, perceptuomotor adaptation (see Section IV,B). A unified approach to adaptation requires us to eschew such disparate conceptual positions. Perception works as an adaptive response because it permits the coordination of action in regard to a real environment, not because it delivers an equivocal central representation of that environment. Similarly, learning works because it permits the development of effectivities that are supported by affordances in a real environment, not because it allows the animal to build an internal model of the world (whether of cognitive or of S–R elements). Section IV consists largely in an elaboration and defense of these assertions.

E. LEARNING AND ADAPTATION—AN OVERVIEW

Before passing on to a detailed consideration of adaptation over the long, short, and medium term, it may be as well to review briefly the vantage point that we have gained thus far with regard to an ecological account of learning.

Our first concern was to give a description of the ecosystem, the fundamental unit of ecological inquiry, and of the two components that comprise it: the actor and its econiche. Actor and econiche were seen to stand in a very intimate relation to one another; indeed, we argued that each must be defined *in terms of* the other. An ecological account of learning, on this view, would appear to be more concerned with relationships between actor and econiche than with either alone. More precisely, we argued that learning should be construed as a process defined over the entire ecosystem, rather than the animal alone in the style of traditional inquiry into learning. It becomes just as important, therefore, to analyze the ecological support for learning that is provided by the econiche as to analyze the biological support provided by the actor.

Given that the relation between actor and econiche is crucial to our understanding of learning, our next step was to consider the nature of this relationship in more detail. Biologists have made use of the concept of adaptation in defining the animal-environment relationship and so an explication of this concept provided us with a useful starting point. Adaptation was analyzed in terms of the attainment of adaptive goals and it was pointed out that the appropriateness of the econiche for the attainment of such goals by the actor is as important an element of the whole process as is the adaptedness of the actor for such attainment. As already discussed, the presentation of Sommerhoff's formal model of adaptation provides us with a powerful logical tool for the analysis of particular instances of adaptation, including learning, and it is that task that we now turn.

IV. Adaptation on Three Time-Scales

The environment of any organism is a dynamic system, characterized by a multitude of ecological factors that change on innumerable time scales. In order to preserve its adaptive relationship with the environment, the organism must be able to adapt concurrently to all relevant scales of change in the environment. Only some such adaptive responses may reasonably be thought to fall in the domain of a theory of learning and in subsequent discussion we shall seek to characterize those responses in ways that are revealing of the kind of theory that will be required to account for them. We are mindful, however, of our broader aim of seeking a theory of learning that can ultimately stand as part of a general theory of the unified adaptive response of an organism to its environment. Our strategy will therefore be to compare and contrast learning with other forms of adaptive response in an attempt to uncover both important similarities and differences between them. Our aim will be, in the spirit of metatheoretical inquiry, to raise illuminating questions about learning that may be answered by some future ecological theory of learning.

A. LONG-TERM ADAPTATION—EVOLUTION BY NATURAL SELECTION

All biological individuals exist as members of more or less extended populations and it is in such populations that evolutionary adaptation is effected. The structure of the gene pool of the population, by which is meant the relative frequencies of different alleles, the distribution of pleiotropic effects, dominance relationships, and so forth, is determined by the nature of the selection pressures that have acted on the population in the course of its evolutionary history. We can perform a thought-experiment in which a genetic population is placed in a situation in which there are no selective pressures acting on it. Under such circumstances the genetic constitution of the population is free to vary in a manner determined by the essentially random factors of mutation and drift. (For the sake of discussion, we will assume the population to be genetically isolated, ignoring the phenomena of migration.) A source of selection pressure, we might say, is some property of the environment that exerts a *constraint* on the free variation in the gene pool. This constraint is exerted because, given the environmental property of interest, some individuals reproduce more successfully than others. Reproductive success depends on the ability to attain the goal of self-reproduction and attaining this goal is, as we have argued (Section III,A), dependent on the ability to attain other adaptive goals that contribute to individual survival and eventual reproduction. Goal-attainment is a function of an individual's phenotype and so differences in reproductive success are attributable to phenotypic differences. A constraint on the free variation in the gene pool of a population, however, can be effected only if certain *genotypes* reproduce more successfully than others. Therefore it follows that natural selection can act in a population only to the extent that differences between genotypes are correlated with phenotypic differences in reproductive success.

In attempting to understand the adaptation of an animal to its environment, we must be concerned not only with its adaptation for reproduction but also with its adaptation for other adaptive goals, such as feeding, locomotion, orientation, and so forth. In order for adaptation in regard to feeding (for example) to be effected by natural selection, there must be a constraint on free variation in the gene pool such that those genotypes whose phenotypes develop the necessary feeding adaptations are reproductively more successful than other genotypes. This can occur only if possession of the adaptation(s) in question is correlated with greater reproductive success. Note that natural selection does not necessarily produce adaptations (Dobzhansky, 1942; Ghiselin, 1966, 1974; Lewontin, 1979; Williams, 1966): *Any* phenotypic characteristic that is correlated with greater reproductive success will be selected, whether or not it contributes to adaptation. Similarly, possession of a phenotypic adaptation is not in itself sufficient to guarantee natural selection in favor of those genotypes that develop such phenotypes: Possession of the adaptation must, in addition, be correlated with greater reproductive success. Our present concern, however, is with those instances in which natural selection does produce adaptation to the environment and with the proper analysis of such events.

We may now identify the elements of Sommerhoff's model of adaptation in the preceding account of natural selection. For clarity, we will employ the following hypothetical, but quite realistic example: At time t_0, a food-limited population gains access to an area containing a new food source not found in its original range and, as a result of natural selection, becomes able to exploit this food source at some later time t_k. First, there is an environmental property (ϕ), namely, those characteristics of the food source relevant to its exploitation by the animal in question, that takes the value ϕ_k at t_k. Second, there is a phenotypic characteristic (β) of a population-typical individual (see footnote 2) that takes the value β_k at t_k. In this example, β might be some aspect of tooth structure and there is some value of β (β_k) that enables an animal to penetrate a hard shell (ϕ) covering the new food source. The goal event G in this example is the ability of a population-typical individual to exploit the food source and G is attained when the focal condition $F(\phi_k, \beta_k) = 0$ is satisfied.

The coenetic variable y_0 in this case, as in most (if not all) cases of evolutionary adaptation, is the same as the environmental variable ϕ to which adaptation is effected. Certain characteristics of the food source (its hard shell, ϕ) constrain free variation in the gene pool because those genotypes whose phenotypes develop values of β equal or close to β_k are reproductively more successful than other genotypes. If the characteristics of the food supply do not change over the course of the population's adaptation,[4] then the mapping function $P(y_0) = \phi_k$ is the identity function (i.e., $y_0 = \phi_k$).

The response of the adapting system, $B(y_0) = \beta_k$, requires more extended analysis than we can provide in this article. As we have seen, the constraint provided by y_0 is on the variation in the gene pool, expressed as differential reproduction among the genotypes that comprise it. But the characteristic β that enables the population-typical individual to exploit the new food supply is measured in the phenotype. The mapping function B is thus composed of two functions. The first of these *(S)* we may call a *selective function,* which maps a constraint, y_0, on to a population-typical genotype (Λ) at time t_k:

$$S(y_0) = \Lambda_k \tag{8}$$

The second function *(E)* is an *epigenetic function* that maps the population-typical genotype on to a population-typical phenotype (charac-

---

[4] This is clearly a rather unrealistic assumption, since the population whose adaptation we are considering will itself exert selection pressure on the prey population, causing the latter to make an adaptive response in turn. These reciprocal effects are, however, too complex to permit their incorporation into this discussion. Nonreciprocal change in the environmental variable ϕ is considered below.

terized by β, among many other features) in the course of individual development at t_k:

$$E(\Lambda_k) = \beta_k \tag{9}$$

The nature of functions such as S is fairly well understood and such functions figure prominently in population genetic theory (e.g., Lewontin, 1974). Functions such as E, on the other hand, are very poorly understood and hardly figure at all in contemporary evolutionary theory. A few authors have attempted to bring developmental considerations into evolultionary theory (e.g., Baldwin, 1902; DeBeer, 1958; Ho & Saunders, 1979; Løvtrup, 1974; Schmalhausen, 1949; Waddington, 1957; see Gould, 1977; Stearns, 1977), but these must be regarded as very preliminary steps. Lewontin (1974, pp. 12–16) provides a brief but insightful assessment of the shortcomings of current theory in this regard. Processes of learning are clearly involved in epigenetic functions and further discussion of this issue will be deferred to Section IV,C.

Finally, we come to the back-reference period, $t_k - t_0$. It will be recalled from previous discussion that the back-reference period is the operation lag of the adaptive mechanism. The precise back-reference period for the process of natural selection will vary depending on a number of factors, in particular the genetic variability of the population and the strength of the selection pressure exerted by y_0. If the genetic variance associated with the phenotypic character β is large, then selection may act rapidly to move the population-typical phenotype from β_0 to β_k. Similarly, if the difference in reproductive fitness between β_0 and β_k is high, producing strong selection pressure, then the population-typical phenotype may shift rapidly.

However rapidly natural selection is able to act, the back-reference period must be at least one generation time, since a change in the population-typical genotype can occur only between successive generations. Generally speaking, natural selection requires much longer than a single generation to effect any adaptive change in a population, often tens or hundreds of generations. The establishment of complex adaptations, such as those required for terrestrial locomotion or flight, may require very much longer periods of time (Frazzetta, 1975). The back-reference period of adaptation by natural selection will, for the sake of discussion, be assumed to be about 10^2 generations, give or take one order of magnitude.

The ability of natural selection to effect adaptation to any feature of the environment is limited by a number of factors. In the preceding example, the adapting population was subject to only one source of selection pressure but in all real situations, populations are subject to a constellation of selection pressures, many of which may tend to operate in opposition to one another. The result is selection for the best available (i.e., reproductively most successful) adaptive compromise and the precision of adaptation that can be made to any one feature of the environment is therefore limited by the adaptive demands of other features. Other limitations include the lack of appropriate genetic variation, the influence of pleiotropic and correlated growth effects, and insufficiently strong selection pressure. We cannot discuss these limitations in detail but they constitute an important and often overlooked aspect of evolutionary adaptation (see Darwin, 1859; Dobzhansky, 1942; Ghiselin, 1966; Gould & Lewontin, 1979; Lewontin, 1979; Williams, 1966).

For our present purposes, the most important limitation of natural selection as a mode of adaptation to the environment lies in its very long back-reference period. The preceding example envisaged a change in the environment from one invariant state (absence of the food source) to another (presence of the food source). Unless other limitations (see above) are in effect, adaptation to this new environmental feature may clearly be achieved through natural selection, since there will be continual selection in favor of phenotypes close to β_k until the population-typical phenotype equals β_k and the focal condition is satisfied. If the environmental variable ϕ changes in the interval $t_0 - t_k$, then it appears that the focal condition $F(\phi_k, \beta_k) = 0$ will never be satisfied and that adaptation cannot be attained until ϕ reaches some stable value.

A moment's reflection shows that one reason for this implausible conclusion is that the focal condition defines too strict a criterion for most real examples of adaptation. In firing a gun at a target (a simple example of goal-directedness) a hit may be scored within some small but finite area around the center of the target. If we designate a hit on the exact center as satisfying the focal condition $F(\phi_k, \beta_k) = 0$, then a hit on the target may still be scored provided the focal condition $F(\phi_k, \beta_k) \pm d = 0$ is satisfied, where d (the *tolerance* of the focal condition) corresponds to a small area close to the target's center. In a similar vein, there will usually be some range of values of an animal's phenotype close to β_k that permits the adaptive goal (i.e., feeding on a new food source) to be attained. This is true for almost all examples of biological adaptation but to simplify notation and discussion we will omit further reference to the tolerance of the focal condition. Tolerance is a quantitative rather than a qualitative addition to the model and its detailed consideration lies outside the scope of our analysis.

We see, then, that natural selection may effect adaptation to environmental variables that are either invariant or that change only slowly in relation to the generation time of the adapting population. Note that the invariance in question may be an invariant *pattern* of change in the environment. Light intensity, for example, changes on a rapid, diurnal cycle but since the pattern of diurnal change remains invariant, natural selection may effect adaptation to this pattern.

Where the relation between the coenetic variable y_0 and the environmental variable ϕ_k is such that the mapping function $P(y_0) = \phi_k$ is indeterminate (corresponding to rapid variation in ϕ over a back-reference period on the order of 10^2 generations), natural selection will be ineffective in producing the requisite adaptation. In each generation there will be selection in regard to ϕ, but the constraint imposed on free genetic variation will change irregularly with the value of ϕ. Over periods of time comparable to a back-reference period of 10^2 generations, there will be only stochastic genetic change in regard to the phenotypic variable β. Adaptation to a rapidly changing environmental variable can be effected only by an adaptive mechanism with a back-reference period comparable to the time-scale of the change. We turn now to consider two such mechanisms.

B. SHORT-TERM ADAPTATION—COORDINATION OF PERCEPTION AND ACTION

Let us first of all consider situations in which the back-reference period of adaptation is brief—on the order of seconds or minutes (see Table I). Such situations comprise almost all of the day-to-day interactions of animals and their environments and so occupy a prominent position in any unified account of adaptation. To illustrate: A person sitting with arms resting on a desk top cluttered with books and papers adjusts the posture of the body in general, and of the arms in particular, to reach for, grasp, and retrieve a book from beneath a pile of papers toward one edge of the desk. Such short-term adaptations have been the focus of considerable analysis in philosophy and psychology; they subsume what are commonly called "perceptions" and "actions."

1. The Phenomenalist Tradition

What we intend in this section is to review briefly (but, we hope, adequately) the epistemological issues to which the study of short-term adaptation is heir. Those issues reduce fundamentally to one question: Are the objects of an animal's perception, with reference to which it behaves, the same as the objects of the animal's environment? The weight of argument over the centuries has tended to be that they are not—that the objects of perception and the objects of the environment are in fact quite distinct. The British philosopher, John Locke, for example, argued that there is an environment that exists independently of the perceiver, who is linked to that environment by means of "ideas"; these ideas, which in some but not all cases represent actual properties of the environment, constitute those things of which the perceiver is directly aware. Locke's "ideas" might be termed "between things" for he intended them as entities that intervene between, or coordinate, the animal and its environment. The traditional generic term for a "between thing" in philosophy is a "phenomenal object," of which some specific contemporary examples are representations, models, reference signals, propositions, and schemata. The term "phenomenalism," therefore, applies to those interpretations of perception in which phenomenal objects, not environmental objects, are what an animal directly experiences and with respect to which it directs its behavior.

Phenomenalist interpretations of perception are of two kinds: those that deny the existence of any but phenomenal objects; and those that admit both phenomenal and environmental objects but claim that only phenomenal objects are involved in the coordination of perception and action. The former view, whose foremost proponents were Berkeley and Hume, has held little attraction for psychologists—it would be a strange science of behavior that attempted to explain adaptation to a nonexistent environment! The latter view, which might be more judiciously termed representative or *indirect realism* (see Cornman, 1975; Mundle, 1971), has, however, been the staple philosophical diet for much of psychology, sensory physiology, and cognitive science. This Lockean view of perception has two major themes: first, that there are environmental objects that exist unperceived and that are unaffected by being perceived (hence "realism"); second, that such environmental objects are not perceived directly but only through the agency of phenomenal objects or "between things" (hence "indirect").

The indirect realist's account of perception distinguishes between what an object *is* (in itself) and what that object *means* (to an animal). A description of what an object *is* is given in conventional physical terms (such as mass, length, velocity, etc.); such a description is not specific to any particular animal and so it is not a description of what the object *means*. An animal behaves with respect to objects in its environment in terms of what they mean for it, however, rather than what they are as crass physical entities, and conventionally it is supposed that the animal ascribes meaning to the physical description of its environment. In other words, the animal interprets the physical description of an object, producing thereby a different kind of (phenomenal) object, describable in terms that are animal-relevant and with respect to which it can behave adaptively.

If we consider the phenomenalist interpretation of perception in terms of Sommerhoff's (1950, 1969) formal model, we see that it yields a most curious account of adaptation over the short term. Under this interpretation, the environmental term, ϕ_{k}, in the focal condition, $F(\phi_{\mathrm{k}}, \beta_{\mathrm{k}}) = 0$, refers to an extraordinary (in the sense of nonreal) property, one that is attributed by the animal to its environment and hence that does not persist unperceived. To illustrate this phenomenalist interpretation, take the case of an animal traversing natural terrain. As the animal encounters obstacles to locomotion and configurations of surfaces that necessitate jumping over, climbing over, or going around, it must adjust its locomotor behavior accordingly. The animal traverses those surface configurations that can support its locomotor activity and skirts those that cannot. It steers through openings that are large enough and around those that are too small. The animal's adaptive acts are with reference to environmental properties such as jump-over-able, walk-on-able, and run-through-able. Yet the time-honored theories of perception, buttressed by the hypostatizing of the basic variables of physics, inform us that such properties are phenomenal rather than real. In short, and this is the larger point, under a phenomenalist interpretation of perception some of the variables over which the focal condition is defined are *not real variables*. The implications of this point for the explanatory scope of contemporary cognitive science have not gone unremarked. Fodor (in press), for one, has argued cogently (and unregretfully) that the current and traditionally popular phenomenalist approach to the knowings of man and animals is a "methodological solipsism" and that questions of what these knowings refer to are beyond its purview.

2. *The Ecological Alternative*

The outcome of the phenomenalist tradition, at all events, is that it leads to a nonunified view of adaptation. The account of long-term, evolutionary adaptation, if it is to be at all sensible, must be given in a vocabulary of real terms on both the animal and environment sides. The account of short-term adaptation, as conventionally construed, is, as we have seen, given only partially in real terms. The conclusion to be drawn, therefore, is that adaptation over the short term is radically different in kind from adaptation over the long term, requiring analysis as a separate phenomenon, *sui generis*. On the phenomenalist view, then, whereas the focal condition of long-term adaptation may be written as $F(\phi_{\mathrm{k}}, \beta_{\mathrm{k}}) = 0$, that of short-term adaptation is to be written as $F(\psi_{\mathrm{k}}, \beta_{\mathrm{k}}) = 0$, where ψ_{k}

is in the codomain of some function $M(\phi_k)$. The function M is construed as a psychological operator that translates meaningless animal-neutral descriptions (such as ϕ_k) into meaningful animal-relevant ones (such as ψ_k). This proliferation of variables is not in itself overly disturbing; what is disturbing is that the additional variable ψ_k is ontologically distinct—it is nonreal, or phenomenal. In order to reconcile adaptation over the long and short terms and to establish a unified account of adaptation, we must eliminate from the focal condition variables of the type ψ_k.

It is not difficult to see that phenomenalism follows in large part from assuming the independence of animal and environment (see Fig. 1) and adopting an interactional style of inquiry. If the animal term and the environment term are logically independent then a third term (the phenomenal object, ψ_k) must be introduced to coordinate the two. The ecological perspective, adopting a transactional style of inquiry, produces a different outcome. As argued at length elsewhere (Shaw & Turvey, in press; Shaw *et al.*, in press; Turvey & Shaw, 1979) and as outlined in Section II,C, the ecological perspective assumes a logical dependence of animal and environment, a dependence that is reflected in the account that we have given of the relation between the actor and its econiche. Because these two terms are mutually dependent, there is no encouragement for a third class of terms to bind them together.

It is here in particular, in eliminating the need for mediation between animal and environment, that the concept of affordance becomes especially significant. As we have said (Section II,A,2), an affordance is some property of an environment taken with reference to an actor; it is thus an animal-relevant property, a component of an econiche, but it is not a phenomenal object. It does not come into and go out of existence with fluctuations in an animal's needs and abilities. An affordance is a *real* property of an environment but it is a part of ecological, not physical, reality (see Gibson, 1977, 1979; Shaw & Turvey, in press; Shaw *et al.*, in press; Turvey & Shaw, 1979). If, then, we describe the environment in animal-relevant terms, we accomplish two things. First, we dispense with the need for mediation between animal and environment by considering instead mutually defined actors and econiches. Second, we preserve the unity of our account of adaptation by defining the focal condition of adaptation over a real environmental variable, an affordance.

3. Describing the Environment—Dimensions of Ecological Physics

In brief summary, casting our account of short-term adaptation in terms of affordances allows us to replace the nonreal variable ψ_k by a real variable, an affordance, ϕ_k. It will be clear from the foregoing discussion that this latter variable is not the same as the physical variable from which ψ_k was derived by the psychological operator M [i.e., by the function $M(\phi_k) = \psi_k$]. To illustrate the distinction, consider the notion of distance, a property that animals perceive and in regard to which they regulate their behavior. In conventional analyses of perception, "distance" is an animal-neutral dimension, measured in some standard, universal metric, such as feet or meters. On such an analysis, the physical distance (ϕ_k) between an animal and an object is translated into the phenomenal variable ψ_k by the function M; ψ_k is what the distance ϕ_k *means* to the animal in the present context (for example, whether the object is reachable and, if so, how much force is needed to propel the body to it). On the ecological analysis, "distance" is an ecosystem dimension, measured in an animal-relevant metric that is defined by the behavioral capabilities, the effectivities, of the actor. Hence the perception of "distance" and the perception of "the behavioral implication of distance" are one and the same.

The nature of distance as an ecological dimension is illustrated by the behavioral relation between a predator and its prey. (This discussion is owing to T. Alley, personal communication, August 1979.) A predator must be able to perceive the maximum distance between itself and a prey animal at which a pursuit can be successfully initiated; and a prey animal must be able to perceive the minimum separation, a "margin of safety," beyond which it need not make defensive or flight maneuvers with reference to a predator. These "distances" are defined, not in reference to an arbitrary metric (such as feet or meters), but in reference to the effectivities of the animals involved and what the current terrain affords them in the way of pursuit and evasive behavior, respectively.

Let us now elaborate this point further with respect to two examples of the coordination of perception and action, basing our analysis on Sommerhoff's formal model of adaptation.

a. Prey Capture by the Praying Mantis. During postlarval growth, the praying mantis goes through several ecdyses (shedding of the exoskeleton), each ecdysis being followed by a rapid growth spurt, after which the exoskeleton hardens again. Each of these developmental stages is called an instar. In each instar, the mantis strikes at prey only within a maximum catching distance (MCD) which bears a definite relation to the maximum extension of the forelimbs (Balderrama & Maldonado, 1973). Striking at prey outside the MCD is presumably wasteful of time and energy and may alert other prey to the mantis' presence, or reveal the mantis to its own predators. In this instance of short-term adaptation, the goal *(G)* is to strike only at catchable prey and for convenience we may assign the following values to ϕ_k and β_k:

ϕ_k = 1 if mantis–prey distance $\leq$ MCD and 0 otherwise;
β_k = 1 if the response is to strike and 0 otherwise.

Then the focal condition may be written as:

$$\phi_k - \beta_k = 0 \qquad (10)$$

and G is attained when this condition is satisfied. Since the MCD changes with each ecdysis, the question arises how the mantis is able repeatedly to attain G throughout development, as it passes through the successive instars. Answering this question amounts to identifying the coenetic variable, y_0, that maps on to an adaptive response β_k, regardless of the current MCD.

Maldonado, Rodriguez, and Balderrama (1974) argue that at each instar, perception of the distance between the mantis and its prey is based on a triangulation system involving three dimensions of the head: the head breadth (HB), the ocular globe breadth (OGB), and the ocular prominence (OP). The coenetic variable is thus some function D of these three variables: $y_0 = D(\mathrm{HB}, \mathrm{OGB}, \mathrm{OP})$. During postlarval growth, these head dimensions maintain a constant relation to the MCD so that *perceptions of distance that are based on them are automatically scaled to the growing animal's behavioral capabilities.* In short, perception of the distance to the prey animal and perception of what that distance means behaviorally are one and the same. This direct adaptive relationship between the mantis' behavior (β_k) and its environment (ϕ_k) is made possible by the involvement of a coenetic variable that describes the environment in terms of the capabilities of the perceiver.

b. Host Tree Location Behavior in a Tropical Vine. Our second example of short-term adaptation is drawn from the plant rather than the animal kingdom. We have selected this example because it demonstrates with particular clarity the importance of analyzing adaptation in terms of the actor and its econiche, rather than in terms of the animal (or plant) alone. It also illustrates the wide range of phenomena that may be embraced by the unified approach to problems of adaptation that we have adopted. Strong and Ray (1975) have described an interesting pattern of behavior in *Monstera gigantea,* an arboreal tropical vine whose seeds germinate on the ground after falling from the parent plant. Immediately following germination, the seedling grows toward the nearest tree and, after making contact, loses its roots as it ascends the trunk. Here we have an instance of adaptation in which the direction of growth (β_k) is adapted to the bearing of the nearest tree (ϕ_k), enabling the plant to contact the trunk *(G)* and complete its life cycle as a mature, arboreal plant. The focal condition is satisfied when the plant grows toward the tree (i.e., when $\phi_k - \beta_k = 0$).

In analyzing this instance of adaptation we seek to characterize the coenetic variable, y_0, and the function, $B(y_0) = \beta_k$, that effects a response adapted to ϕ_k. Strong and Ray (1975) demonstrated experimentally that the seedling's behavior is an example of positive skototropism, a positive growth response toward darkness: A seedling always grows toward the darkest sector of its horizon. The nature of the response function B remains to be elucidated but, by analogy with other instances of plant tropisms (Bell, 1959), we might suppose that it involves the differential transport of auxins (plant growth hormones) to or from various points on the circumference of the seedling.

In the plant's natural environment, a physical description of the environment would define ϕ_k as "bearing of nearest tree." However, adaptation is effected not in relation to the tree as a physical (i.e., organism-neutral) object but rather as an object that affords climbing for the plant. Adaptation could not be effected if we were to populate the plant's environment with objects that, while conforming to the *physical* description of a tree, did not conform (perhaps because of their surface properties) to the *ecological* description of a climbable object (climbable, that is, by the plant). It is only by carrying out the analysis in respect to the integrated ecosystem, in which actor and econiche are inseparable and defined in terms of each other, that the adaptiveness of this behavior can be understood; and this requires that our descriptions be given in ecological, not physical dimensions.

An understanding of the adaptiveness of the response to the coenetic variable likewise requires that we adopt an ecological description of the plant's environment. Under a physical description of the environment, the darkest sector on the plant's horizon yields "bearing of lowest light intensity" and light intensities do not, of course, support climbing. Under an ecological description, however, the darkest sector yields "bearing of nearest tree" (more precisely, "nearest climbable object") which does support climbing. In the terms of Sommerhoff's model, the ecological description may be given as $P(y_0) = \phi_k$, P being an ecological function, mapping dark sectors on to climbable objects. Once again, we could populate the plant's environment with objects whose properties were such that dark sectors no longer mapped on to climbable objects. Strong and Ray (1975) accomplished this by using the open ends of opaque tubes to produce dark sectors on the plant's horizon. In this situation, the physical description of the environment as an array of light intensities is preserved but the ecological description has been changed because the mapping function P no longer holds. Not surprisingly, the plant's response is now found to be nonadaptive and it grows into the open end of the nearest tube. (This response, incidentally, provides an interesting demonstration of an optical illusion in this species.) We see, then, that the adaptive relationship between the plant's behavior and its environment can be discerned only by adopting an ecological scale of description and respecting the mutual dependence of the actor and its econiche.

4. *The Adaptive Response to Environmental Structure*

It has become common practice, in many contemporary discussions of short-term, perceptuomotor adaptation, to speak of a plan or program that controls an organism's behavior with respect to perceptually delivered information about the environment (see Miller, Galanter, & Pribram, 1960; Turvey, 1977a). Such an entity is one variety of phenomenal object that mediates between the animal and its environment on the output rather than the input side of perceptuomotor adaptation. Like the phenomenal objects of perception, the motor program threatens the unity of our account of adaptation by its implication of additional nonreal variables in the focal condition of adaptation. In conventional terms, a program is a nonreal, phenomenal object created by the organism to stand in an adaptive relationship with the environment and so permit the adaptive control of behavior.

The status of a program in the phenomenalist account of behavior is that of an explicit, a priori description of the orderliness or adaptiveness of behavior. There are two possible ways to view the process by which this description might serve to control behavior. One is to view it as a recipe that is followed by some executive component of the system that directly controls the animal's behavior, in much the same way that one might follow a set of instructions for building a boat. The obvious drawback to this view is that it replaces one problem (accounting for the animal's ability to behave adaptively) with another (accounting for the executive's ability to follow instructions) that inherits all of the logical and psychological problems of the first and so initiates the first step in an infinite regress. The second and less objectionable way is to view the program as being intrinsic to the structure of the behaving animal and as being implicitly rather than explicitly followed.

Cummins (1977) has pointed out that this strategy, which is adopted by most proponents of the metaphor of a motor program or action plan, leaves little or no room for distinction between the structure of the program and that of the system that, by appearances, is executing it. He points out that, on this view, the program is simply a description of those aspects of the structure of the animal that enable it to behave adaptively.

This argument eliminates the program as an explanatory entity, as an a priori prescription *for* the system, and reconstrues it as an a posteriori description *of* the system.

The structure of the animal alone, however, cannot guarantee the adaptiveness of behavior for, as we saw earlier (Section III,A), adaptation is a *relation* between the actor and its econiche. To account for the adaptiveness of behavior, we must therefore describe the animal's structure in terms of its environment. Such a description is provided by the response function in Sommerhoff's model, which maps the coenetic variable of adaptation on to a behavioral variable: $B(y_0) = \beta_k$. The form of this function is defined by the constitution or structure of the animal and it provides a description of that structure with respect to the environment, that is, with respect to the coenetic variable, y_0.

Let us then construe the concept of a "program" as an environment-relevant description of the animal that is provided by the function $B(y_0) = \beta_k$. On this account, perceptuomotor adaptation does not involve the coordination of the animal and its environment through the agency of phenomenal objects. Rather, it involves a *direct adaptive response* to environmental structure (y_0) that constrains the animal's behavior to some particular response, β_k; this constraint is adaptive when the focal condition, $F(\phi_k, \beta_k) = 0$, is satisfied. In the case of the tropical vine *Monstera gigantea,* the directness of the adaptive response is revealed with particular clarity, because of the simplicity of the adaptive system. The plant's perception of its environment (in particular, of the bearing of the nearest tree) is based, let us say, on the differential transport of auxins around the circumference of the stem. But this is also the means by which the plant effects an adaptive response to the layout of the environment, growing in the direction of the nearest tree. We may say, therefore, that the short-term adaptive response is effected by a constraint, specific to the layout of the environment, on the plant's behavior (direction of growth) and that this constraint arises as a direct response to perceived environmental structure.

In the case of more complex organisms, the directness of the adaptive response is obscured by the elaborate physiological support required for its implementation. The response function nonetheless provides a formal, albeit abbreviated description of those aspects of the organism's structure whose sensitivity to the coenetic variable (y_0) produces a response (β_k) over the back-reference period of adaptation. Let us adopt the term "perception/action system" to refer to the relevant aspects of organismic structure. [For further discussion of this and related concepts, see Fitch and Turvey (1979), Fowler (1977), Johnston (1978), Turvey (1977a), and Turvey, Shaw, and Mace (1978).] The perception/action system may be in any of an indefinite number of states and the state of the system determines the current behavioral performance (β_k) of the organism (e.g., striking or not striking at a prey, or growing in any one of an indefinite number of directions). The value of β_k thus reflects a constraint on the state of the system that is specific to the coenetic variable y_0 and that is adaptive when the focal condition is satisfied.

While the *state* of the system (specific to β_k) is relatively transient, being specific to the organism's current perception of its environment, the *structure* of the system, as described by the response function, $B(y_0) = \beta_k$, is relatively more enduring. This structure arises in the course of individual development and it is here, in the epigenetic processes that constrain the course of development, that we encounter the subject matter for the study of learning.

C. MEDIUM-TERM ADAPTATION—LEARNING

In preparation for our discussion of learning as a form of adaptation over the medium term, let us briefly review the story we have told thus far of the nature of adaptive response to the environment. On both the long and the short terms, we spoke of constraints, specific to certain features of the environment, arising in the course of adaptation. In the case of evolutionary adaptation over the long term, constraints arise on free genetic variation in the gene pool that are specific to slowly changing features of the environment, essentially invariant over periods of time comparable to the life span of an individual. In the case of perceptuomotor adaptation over the short term, the constraints are specific to much more rapidly varying features of the environment, in particular to the affordances whose availability changes over time as the animal moves about.

On both the long and the short term, constraints arise as a direct adaptive response to environmental structure. In evolutionary adaptation, they do so by virtue of a sensitivity of the gene pool to sustained selection pressure, a sensitivity that is expressed as the differential reproduction of genotypes. In perceptuomotor adaptation, they arise by virtue of a sensitivity of the perception/action system to information in the form of structured energy (such as light and sound), a sensitivity that is expressed as a modification of the state of the system, hence in the form of the actions that it specifies. It is neither remarkable nor problematical to speak of the adaptive response of the gene pool as direct. No theory of indirect evolutionary adaptation (in which a representation of the environment is constructed, the adaptive response being specific to that representation) has ever been proposed; indeed, it is hard to see what might be meant by an indirect response in the context of evolutionary adaptation.

Theories of indirect perception, on the other hand, are commonplace in psychology but we have argued for a theory of direct perception, a position commensurate with our earlier arguments (Section III,D) in favor of a unified theory of biological adaptation.

Our approach is consonant with a philosophical position that we have called ‘‘pragmatic realism,’’ and this position will guide our inquiry into the medium-term adaptation provided by learning. Particular environments place particular adaptive demands on the animals that live in them. Natural selection ensures that those individuals that effect pragmatically successful responses to those demands, responses that ensure survival and eventual reproduction, will come to be typical of the population. Insofar as such responses entail adaptation to certain rapidly varying aspects of the environment, then perceptual abilities will evolve that are specific to those aspects of the *particular* environment in which the population is evolving.

Standing between the long-term constraints on the gene pool effected by natural selection and the short-term constraints on the state of the action system effected by perception is a set of constraints that arise as a result of relatively prolonged epigenetic processes. Some of these medium-term constraints reflect adaptive responses of the kind that we would wish to identify as learning and the account of learning that we will give parallels those already given for adaptation on both shorter and longer time-scales. We will argue that learning, like the other forms of adaptation we have been discussing, is a direct adaptive response, in this case to aspects of the environment that change over periods of time that are short in comparison with evolutionary time-scales but long in comparison with the events of perception. Furthermore, the learning abilities that are typical of a population are those that are ‘‘pragmatically successful’’ in the particular environment in which the population evolves. Our account of learning will be guided by these two principles of direct adaptation and pragmatic realism.

An ecological account of learning as a direct adaptive response to a particular environment faces two central problems. First, it must provide an appropriate description of the environment that is being adapted to. In the terms of Sommerhoff's (1950) model, this means identifying the environmental variable (ϕ) to which adaptation is effected, the coenetic variable (y), and the mapping function [$P(y_0) = \phi_k$] that relates the two over the back-reference period of adaptation, $t_k - t_0$. Second, it must provide an account of the adaptive response that the animal makes to the coenetic variable, identifying the phenotypic basis of adaptation (β) and defining the mapping function [$B(y_0) = \beta_k$] that is implemented by the adaptive response.

1. Describing the Environmental Support for Learning

The current, nonecological approaches to the study of animal learning adopt a very different attitude toward the problem of environmental description than does the ecological approach. In the associationist account of learning (which still largely dominates current thinking in the field; see Jenkins, 1979) the environment is seen as an array of stimuli and, in some versions, of reinforcers. The concept of stimulus in such theories is entirely nonecological and its definition is not specific to any particular organism. Anything to which an animal can be persuaded to respond by an experimenter counts as a stimulus in association theory. An ecological approach to learning, however, must treat the problem of environmental description quite differently and indeed must accept it as a significant component of the overall research endeavour.

In discussing the structure of the ecosystem in Section II,B, we argued that the econiche is a description of the environment taken with respect to some actor. Specifically, it is a description of the ecological support for behavior and in the preceding discussion of perception we showed how this support may be exploited by a suitably attuned actor. In the same manner, we now argue that an ecological approach to learning must begin with a description of the ecological support that an environment provides for a suitably attuned learner. Such a description cannot be phrased in animal-neutral terms. Animals become attuned, in the course of evolution, to particular aspects of environmental structure that support learning. They evolve the particular attunements (i.e., learning abilities) that they do because such adaptations are pragmatically successful in the environment in which the population has evolved. To the extent that relevant aspects of environmental structure are unique to particular ecosystems, then we expect to find specialized learning abilities, limited to one or a few species. To the extent that these aspects of structure characterize many ecosystems, then we may expect to find large numbers of species attuned to them. Any animal's learning abilities, however, are collectively a complex adaptation to *particular* aspects of environmental structure, namely, those that characterize its particular ecosystem, whether these are of widespread or restricted occurrence. An ecological account of learning requires a description of environmental structure as an integral part.

The ecological support for learning referred to in the preceding paragraph is, of course, the coenetic variable of adaptation (Sommerhoff, 1950). It is that aspect of the environment to which the adaptive response is made and that therefore provides the ecological support for the learning ability in question. In many cases of learning the coenetic variable will be the same as the environmental variable (ϕ) to which adaptation is effected as a result of learning. This will not necessarily be the case however and in Section V we discuss some instances of learning in which y and ϕ are different variables; that is, adaptation to one aspect of the environment (ϕ) is effected by virtue of a developmental sensitivity to some other aspect (y).

It is perhaps worth stating explicitly that the description of y and ϕ is an empirical problem that can be solved only by studying particular environments in relation to the particular animals that live in them. No general answer can be given to the question "What constitutes the environmental support for learning?" This question must be posed separately for each instance of medium-term adaptation in a specific ecosystem, and answered on the basis of empirical investigation. We can, however, offer some general considerations to guide such investigation.

In our discussion of evolutionary adaptation in Section IV,A, we pointed out that long-term adaptive responses are effective only in regard to slowly changing features of the environment. Suppose that we were interested in the evolutionary response of a population to change in environmental temperature. We can describe change in temperature at any number of "grains of analysis" (Fitch & Turvey, 1979), from the very small, moment-to-moment changes that are produced by wind currents and shadows, to the very gradual shifts in mean annual or decadal temperature that are produced by climatic changes operating over continental expanses. If, in searching for the coenetic variable of evolutionary adaptation to temperature (that is, the ecological support for this form of adaptive response), we focused on the microclimatic changes in temperature to be found at a very fine grain of analysis of the environment, one of two situations might arise. In the first place, we might fail altogether to detect the gradual changes in temperature that support evolutionary adaptation—a case of being unable to see the forest for the trees. On the other hand, we might detect the gradual changes by recording very many small changes in temperature and integrating these over long periods of time. In the latter case, we would then be faced with the problem of explaining how the adapting population performs an analogous integrating operation, identifying the nature and location of the computational machinery, and how it produces an adaptive response based on the outcome of this integration. Evidently, this is an artificial problem that arises because we have adopted an inappropriately fine grain of analysis of the environment. It is only when we provide a description at a coarser grain of analysis, appropriate to the long-term nature of evolutionary change, that we detect the gradual shifts in temperature that provide the ecological support for a direct adaptive response, unmediated by any form of integration.

In searching for the ecological support for medium-term processes of adaptation such as learning, we must adopt a similar strategy of focusing at an appropriate, intermediate grain of analysis. This same strategy has been proposed by Humphrey (1933) in a remarkable and neglected book, *On the Nature of Learning.* Humphrey pointed out that an animal may respond adaptively to a wide range of environmental events (or aspects of structure), some of which are of very brief duration, others much more prolonged. He suggested, as we have done, that the attempt should be made to account for all such adaptive responses under the same theoretical rubric, rather than treating each one as a phenomenon *sui generis.* In pursuit of this aim, we must describe the ecological support for each form of adaptation at an appropriately fine or coarse grain of analysis. If our description of the environment is provided at too fine a grain, then we may either fail to detect the ecological support for learning, or we may be forced to postulate constructs such as memory and associations to mediate between the apparently unconnected elements of our description. If, on the other hand, we seek a coarser grain of description of the environment, one that is more appropriate to the medium-term adaptive response of learning, the need for such constructs may well not arise. In that case, we shall have uncovered the ecological support for learning as a direct adaptive response. In Section V, we will discuss some examples of learning that meet these expectations.

Note the similarity between this orientation to the problem of learning and that of Gibson (1966) toward the problem of perception. Gibson's program was initiated with the claim (Gibson, 1950) that far greater environmental support for perception could be discovered than had previously seemed to be the case. Rather than postulating epistemic mediation (Turvey, 1977b) to account for the elaboration of impoverished sensory data into richly structured perceptual experience, Gibson claimed that the environment itself provides a source of richly structured stimulative energy, to which an appropriately attuned perceiver might respond. Uncovering this structure, however, requires that an appropriate style of description be employed, one that focuses at a suitably coarse grain of analysis (Fitch & Turvey, 1979). We claim, together with Humphrey (1933), that the need for epistemic mediation in learning might likewise be eliminated if we were but to employ appropriate descriptions of the situations in which such adaptive responses occur.

2. *The Adaptive Response to Environmental Structure*

In Section III,A we pointed out that the response function involved in evolutionary adaptation [$B(y_0) = \beta_k$] may be resolved into two functions: a selective function [$S(y_0) = \Lambda_k$], which maps the coenetic variable (selective pressure) at t_0 on to a population-typical genotype at t_k; and an epigenetic function $[E(\Lambda_k) = \beta_k]$, which maps the genotype on to a population-typical phenotype. In that discussion, the epigenetic function was implicitly assumed to operate instantaneously at t_k. In fact, this function is itself composed of a number of adaptive processes, with back-reference periods covering the life span of the individual, some of which fall in the domain of an ecological theory of learning. In this section we will consider the nature of the medium-term adaptive responses that implement the epigenetic function. We will continue to use the notation of Sommerhoff's model and the reader should note that, for the remainder of this section, the variables of this notation (y_0, t_0, t_k, β_k, etc.) will apply to the medium-term adaptive responses of learning, unless explicit mention to the contrary is made.

The epigenetic view of development is widespread among modern students of development and is associated especially with the names of Kuo (1967), Lehrman (1953), Schneirla (1956, 1965), and Waddington (1957). Our presentation will be brief and is intended to provide a basis for our account of learning, not as a comprehensive treatment of the issues. The developing organism is subject to two sets of constraints that mutually determine its organization; one of these sets originates in the genome, the other in the environment. The genetic constraints specify what we might call a "life-strategy" for the individual, to which the environmental constraints provide a set of tactical modifications. In some respects, the genetic strategy may be quite strictly defined, providing what Waddington (1957) calls strongly canalized development, relatively insensitive to the nature of the environment. In other respects, the genetic strategy may be less tightly constraining, permitting the nature of the individual's experience to play an important role in determining the course of development. In the latter case, Waddington (1957) speaks of weakly canalized development, the extent of canalization being, of course, a matter of degree.

The development of a particular phenotypic character may, if it is relatively weakly canalized, be sensitive to only a very narrow range of environmental input. For example, isolating the young of many species of songbirds early in life results in a form of adult song that is quite differently structured from the song of conspecifics that have received exposure to an adult song model as youngsters (Marler & Mundinger, 1971; Nottebohm, 1970). In some species, development of normal song organization results only from exposure to conspecific song (i.e., only the species-typical song will be learned), whereas in other species a much wider range of songs will be accepted as song models. Thus while song development in many species is relatively weakly canalized (since song and no-song experience lead to different phenotypic outcomes in adulthood), the selectivity of the developmental response may be much greater in some cases than in others.

Insofar as the question of the selectivity of developmental responsiveness is concerned only with the degree of potential developmental plasticity, it is not an issue of primary concern for the ecological study of learning (see Gottlieb, 1976, for an expression of a similar point of view). We are more concerned with the individual's adaptive response to the *typical* environment of development, although its response to atypical, experimental environments (as in the selective deprivation or selective exposure experiment) may clearly make an important contribution to our understanding of the former. The question of primary interest as regards plasticity is therefore whether an animal will respond to any of a range of experiences (i.e., values of y_0) *typically* encountered in development, or to only one or a few of these experiences. This emphasis follows directly from our identification of the ecosystem (i.e., an actor and its econiche) as the unit of inquiry, rather than the animal, considered in isolation.

Let us suppose, then, that under normal (i.e., species-typical) circumstances, the individual develops some phenotypic character (β) such as the species-typical song of an adult songbird. At some particular time in development (t_k) the organization of the character (β_k) is such that a focal condition $[F(\phi_k, \beta_k) = 0]$ is satisfied and some adaptive goal *(G)*, such as the acquisition of a mate by a vocally competent male bird, can be attained. In this example, ϕ_k might be the tendency of conspecific females to mate with a singing male. If this is an example of medium-term adaptation, then it will be possible to identify a coenetic variable (y_0) and a response function $[B(y_0) = \beta_k]$ that maps the coenetic variable on to the phenotypic character at t_k. In the example of song development, the coenetic variable is exposure, early in life, to an adult song model.

The function B covers a multitude of important issues concerning the mechanisms of adaptation that cannot be properly addressed in this article; we will, however, discuss some of them briefly. First, many instances of medium-term adaptation can be effected only, or best, during a restricted portion of the individual's lifetime, generally in early life. For example, white-crowned sparrows *(Zonotrichia leucophrys)* will learn the characteristics of an adult song to which they are exposed during the first few months of life but not those of songs that they hear thereafter (Konishi & Nottebohm, 1969). In some cases, therefore, it is necessary to assign a range of permissible values to t_0, defining a sensitive period within the life cycle when adaptation can be effected.

Second, the response to the coenetic variable may be such as to affect the course of development in any of several possible ways. Gottlieb

(1976) has defined three roles that experience may play in the development of behavior: maintenance, in which experience is required for the continued development of a preexisting characteristic; facilitation, in which experience acts to speed up development; and induction, in which absence of experience precludes development of the characteristic altogether. In its present form, the model we have employed offers no way of distinguishing among these three roles of experience (and others that might be proposed). Elaboration of the model in this regard will require attention to the relationship between β_0 and β_k and to the nature of the continuous transformation in β over the course of the back-reference period.

Finally, there is the problem of deciding which instances of medium-term adaptation count as examples of learning and hence form part of the subject matter for an ecological theory of learning. At the outset, we would caution against the temptation to delineate hard and fast categories of "learning" and "not-learning." In Section III,D we pointed out that the adaptive response of the animal to its environment is a *unified* response. Particular "kinds" of adaptive response, such as those we have identified in this article, are probably best regarded as modal points on a fairly densely populated continuum of response. If the unified approach that we advocate to problems of adaptation is indeed appropriate, then attempts to make sharp delineations between categories are bound to be theoretically unproductive. On the other hand, we would certainly accept the heuristic value of recognizing paradigmatic or typical examples of what we have called short-term and medium-term adaptive reponses, and using the analysis of such examples to sharpen our appreciation of both similarities and differences among these various forms of response.

Let us briefly consider, then, one or two criteria that seem particularly important in determining the typicality of putative examples of learning. In the first place, we would recognize that the study of learning is concerned with change in an animal's behavior rather than in other aspects of its organization such as its morphology or biochemical make-up. Typical instances of learning are therefore those in which the outcome of medium-term adaptation (β_K) is some measure of the behavioral phenotype. It should be emphasized, as pointed out in Section II,A, that this requirement specifies a choice of a particular *description* of the phenotype; all instances of learning presumably have some physicochemical basis and so may in principle be described in morphological (or physiological or biochemical) as well as in behavioral terms. The primary concern for a theory of learning, however, is to account for the behavioral description of an adaptive change and typical instances of learning will be those in which the behavioral change is regular, consistent, and demonstrably adaptive, rather than being secondary to some other, say physiological, change. Accounting for the physiological basis of learning is, of course, a different matter, with which we cannot attempt to deal here.

A second important criterion concerns the specificity of the relationship between y_0 and β_k. There are many effects of experience that influence an animal's behavior in ways that are indeed adaptive but that are also highly nonspecific. To give but one example, adequate nutrition is well known to be required for proper behavioral development in many species (Leathwood, 1978) but the extreme nonspecificity of such effects precludes their acceptance as typical or illuminating instances of learning. We would not, however, wish this criterion to be interpreted as license to erect two mutually exclusive categories of developmental response to the environment, one specific, the other nonspecific. Rather, we suggest that there is a continuum in the specificity of response, with the more typical instances of learning being located toward the ''most specific'' rather than the ''least specific'' end (see Bateson, 1976, for further discussion). Again, we urge that an ecological approach to learning respect the unity of the adaptive response to the environment and be prepared, at least in its early stages, to embrace a fairly wide range of developmental phenomena in the search for general theoretical principles. For example, the effects of enriched experience on behavioral development (Rosenzweig & Bennett, 1978) seem to lie toward the middle of the continuum just described and are the kind of nontraditional, borderline effects with which a unified, ecological approach to learning should be prepared to deal.

V. Learning as Medium-Term Adaptation—Analysis of Three Examples

Having given an account of some of the conceptual issues involved in an ecological approach to learning as medium-term adaptation, let us now turn to consider some concrete examples of such an approach. In selecting examples of learning to present in this section we have restricted ourselves to cases in which the phenotypic characteristic whose development is being analyzed is of clear adaptive significance to the animal that possesses it. Many of the learning tasks that are employed in conventional studies of learning are of questionable adaptive relevance to the animal outside the artificial setting of the laboratory. Their contribution to the understanding of learning as a form of biological adaptation is therefore not readily apparent. This is not to say that we see laboratory investigation as irrelevant to the ecological study of learning. On the contrary, nothing can be learned about the identity of coenetic variables and the ways in which they influence development without perturbing the natural course of development in some manner. The use of artificial or atypical rearing environments must obviously play a major analytic role in any study of learning, whether ecological or otherwise. But in the ecological approach, such experimental intervention must be designed to analyze an ecosystem rather than an animal (see Section II), seeking to explicate the ecological support for learning and the means by which such adaptation is effected.

An ecological analysis of learning proceeds from the observation that at some point in an animal's life (t_k), some aspect of the animal's phenotype (β_k) and some aspect of its environment (ϕ_k) stand in a particular relationship to one another [i.e. $F(\phi_k, \beta_k) = 0$], such that some adaptive goal (G) may be attained. Analysis of the medium-term adaptation that gives rise to β_k involves identifying the coenetic variable (y) that provides the ecological support for adaptation and describing the function [$B(y_0) = \beta_k$] that produces the characteristic in the course of normal development. The response function B has evolved in an ecosystem in which there is a *specific* relationship between y and ϕ [namely, $P(y_0) = \phi_k$] and between ϕ and β [namely, $F(\phi_k, \beta_k) = 0$]. Analysis of B cannot therefore be based on an arbitrary selection of y, ϕ, and β; selection of these variables must reflect the structure of the particular ecosystem under analysis.

One example of an ecological approach to learning, the development of bird song, has already been discussed, in Section IV,C,2. Three more examples will now be considered.

A. SPECIES IDENTIFICATION IN DUCKLINGS

The young of many species of precocial birds show a strong and specific tendency to approach a species-typical call, the maternal assembly call, which is uttered by the mother to lead the young off the nest after hatching and to ensure cohesion of the brood following nest exodus (Collias & Collias, 1956; Gottlieb, 1965). Thus, at about the time of nest exodous (t_k), the adaptive goal of brood cohesion (G) is attained by virtue of a tendency of the young (β_k) to approach the maternal assembly call (ϕ_k). The specificity of the approach tendency is such that only the call of the young's own species will be approached, and this specificity is defined by a focal condition,[5] $F(\phi_k, \beta_k) = 0$. An analysis of the medium-term adaptation underlying this state of affairs involves identifying the

---

[5] In this and the following examples we will use the formalism provided by Sommerhoff's model without specifying the nature of the function that defines the focal condition. In principle, the necessary specification can always be provided, but it will depend on a number of considerations that

coenetic variable (y_0) and describing the function $B(y_0) = \beta_k$ that effects the adaptive response.

If mallard *(Anas platyrhynchos)* or Peking (a highly domesticated form of mallard) ducklings are raised in an incubator, with no exposure to the maternal assembly call, they will nonetheless show a strong and specific tendency to approach the call after hatching (Gottlieb, 1971). The mallard call is preferred to those of a variety of other species, including the pintail *(Anas acuta),* the wood duck *(Aix sponsa),* and the domestic chicken *(Gallus gallus).* The fact that prior exposure to the maternal call is not required for the development of a specific approach tendency indicates that the coenetic variable of adaptation is in this case not the same as the variable (the maternal call) to which adaptation is effected (see Section III,B).

Identification of the coenetic variable has been achieved on the basis of experiments in which the ducklings' prenatal experience was artificially modified. During embryonic development, the duckling begins to vocalize approximately 3 days before hatching, when it moves into the airspace at the large end of the egg (Gottlieb & Vandenbergh, 1968). If the embryo is surgically devocalized before this time and reared in auditory isolation, so that it can hear neither its own vocalizations nor those of siblings, then a less specific tendency to approach the mallard maternal call will be evident when the duckling is tested postnatally. Specifically, devocal ducklings approach the chicken call about as often as the mallard call in a choice between the two (Gottlieb, 1971). Experiments using artificially altered mallard calls in choice tests revealed that devocal ducklings are relatively insensitive to two acoustic features that differentiate the mallard and chicken calls: a high-frequency component that is lacking from the chicken call (Gottlieb, 1975a) and the repetition rate of the individual notes that make up the call (Gottlieb, 1978).

The specificity of approach shown by normal ducklings to mallard calls containing the high-frequency component may be reinstated in devocal ducklings by exposing the embryo, after devocalization, to a recording of the contact-contentment call (Gottlieb, 1975b), one of the calls that a normal embryo uttters during the last 3 prenatal days. This result clearly identifies exposure to the contact-contentment call as the coenetic variable in this instance of medium-term adaptation. There is considerable specificity between y_0 and β_k in this example, as shown by the fact that exposure to either recordings of distress calls (also uttered by the embryo)

---

lie outside the scope of this article, particularly the choice of suitable measures for β and ϕ. The example of prey capture by the praying mantis discussed in Section IV,B shows one strategy for defining the focal condition that could be employed in any of the examples discussed here.

or to suitably pulsed bursts of white noise is not effective in reinstating the specific approach tendency in postnatal tests (Gottlieb, 1975b).

The tendency of normal ducklings to approach calls pulsed only at the species-typical rate of about four notes/sec may also be reinstated in devocal ducklings by exposing the embryos to the contact-contentment call (Gottlieb, in press). The embryonic call also has a repetition rate of about four notes/sec and the specificity of the developmental response to it is demonstrated by the fact that exposure to the same call pulsed at either 2 or 5.8 notes/sec does not reinstate the specific postnatal approach tendency (Gottlieb, in press).

Gottlieb's research has made it possible to give a very complete account of both the coenetic variable and the function that defines the adaptive response in this instance of adaptation. The former has already been sufficiently discussed; the nature of the response to the coenetic variable is shown by other results reported by Gottlieb. The sensitivity to the high-frequency component of the mallard call, which, in devocal ducklings exposed to the contact-contentment call, appears at about 24 hr after hatching, emerges even in the absence of such exposure by 48 hr (Gottlieb, 1975c), showing that experience plays a facilitative role in its development (Gottlieb, 1976). The role of experience in the development of the repetition rate preference, on the other hand, appears to be one of maintenance, since the embryo exhibits a specific motor response to calls pulsed at the species-typical rate of four notes/sec even before the normal onset of vocalization (Gottlieb, 1979).

B. STELLAR ORIENTATION IN THE INDIGO BUNTING

Many species of birds migrate over considerable distances between their summer and winter ranges, an accomplishment that depends on navigational skills of a high order of refinement. In general, these skills require that a migrant's direction of movement be controlled with respect to some source of information in the environment that specifies the direction of the terminus of migration with respect to the migrant's present position (Emlen, 1975a). Thus, by the time of its first migration season (t_k) the migrant's navigational skill (β_k) is such that it can utilize the available sources of directional information (ϕ_k) to navigate with the required accuracy *(G)*. The focal condition $[F(\phi_k, \beta_k) = 0]$ is satisfied when the migrant can select a seasonally appropriate direction in relation to the available ecological information.

The indigo bunting *(Passerina cyanea)* is a nocturnal migrant that uses the information present in star patterns as the basis for its navigation (Emlen, 1967a). From experiments using projected star fields in a planetarium, Emlen (1967b) concluded that configurational information in groups of stars (rather than the bearings of individual stars) provides the basis for navigation. Since blocking out arbitrary portions of the star field does not affect the birds' directional preferences, the relevant information must be widely distributed over the sky, rather than being inherent in particular groups of stars.

Birds that are denied exposure to the night sky until the start of their first migration season do not show the orientation preferences of normally experienced birds (Emlen, 1969). Acquisition of this skill is therefore an instance of medium-term adaptation, information present in the night sky being identified as the coenetic variable. One month of exposure just prior to the migration season is sufficient to establish a weak, although appropriate directional preference. Further experiments (Emlen, 1970, 1972) showed that definition of the axis of rotation of the sky is required to specify the coenetic variable fully. In order for a bird to orient in a particular direction appropriate to the season, rather than just in a constant arbitrary direction, it must have a directional reference and the axis of rotation of the sky provides such a reference, pointing north in the northern hemisphere. Emlen (1970) reared two groups of birds, one exposed to an artificial (planetarium) sky rotating normally about Polaris (the pole star), the other exposed to the same sky rotating about Betelgeuse. When tested for directional preferences in their first migration season, birds in the first group showed normal directional preferences with respect to Polaris whereas those in the second group showed equivalent preferences with respect to Betelgeuse. Subsequent reexposure of the second group to a normally rotating sky failed to modify their preferences, suggesting the existence of a sensitive period for the acquisition of the navigational skill.

Emlen's analysis of the medium-term adaptation underlying the navigational ability (β_k) of adult indigo buntings has identified the environmental variable (ϕ_k) to which adaptation is effected (configurational information in star groups) and the coenetic variable (y_0) that supports the adaptive response. In contrast to the example discussed previously, ϕ and y refer in this example to the same aspect of the environment (the star field). However, specification of ϕ and y requires different descriptions of the star field; specifying y requires that the axis of rotation be defined, whereas adult birds will orient correctly under a stationary sky (ϕ) (Emlen, 1967a). Note also that whereas in the preceding example only one value of y_0 (a contact-contentment call pulsed at the species-typical rate) would support adaptation, in this example there is some range (S_0) of values (i.e., range of rotational axes), to each of which an appropriate adaptive response is made. Birds raised under skies with different axes of rotation orient differently with respect to configurational information in the stars but all choose an equivalent direction with respect to "celestial north," as defined by the axis of rotation.

As mentioned above, Emlen's (1970) results hint at the existence of a sensitive period in this instance of adaptation. They also suggest that the role of experience is inductive (Gottlieb, 1976) rather than maintaining or facilitative. Inexperienced birds show no directional preference and specific preferences may be induced with respect to any of some range of axes of rotation. This research thus provides at least a partial characterization of the function *B* that effects the adaptive response.

C. VISUALLY GUIDED REACHING IN MAMMALS

Optical information provides perhaps the most prominent source of control on behavior, as demonstrated in experiments by Lee (1976; Lee & Aronson, 1974; Lee & Lishman, 1977), Schiff (1965), Warren (1976), and others (see Turvey, 1977a; Fitch & Turvey, 1979). Consider a simple example of visuomotor control such as reaching for an object in the visual field. In Section III,B we analyzed a similar example as an instance of short-term (i.e., perceptuomotor) adaptation; here we are concerned with the development of the skill itself as an instance of medium-term adaptation rather than with the short-term analysis of its components. Gibson (1958) proposed that an animal might control its behavior by responding appropriately to the patterns of texture gradients and texture flow fields that are present in the optic array (Gibson, 1950), that richly structured optic medium made available by the multiply reflected light in a complex natural environment. Mathematical analyses by Johansson (1974), Lee (1974), and Nakayama and Loomis (1974) have confirmed Gibson's intuition that specificity does indeed exist between patterns, both static and dynamic, of optic texture and the layout of objects and surfaces in the environment. It is significant that in Lee's (1974) analysis, the optical patterns of relevance are scaled to the perceiver's body size; for example, height of the eye above the ground appears as a variable in the equations describing patterns specific to the control of locomotion.

Taken together, the theoretical and experimental results permit the following assessment of the situation at some time t_k in an animal's life: As a skilled perceiver, the animal has the ability (β_k) to use ecological (i.e., body-scaled) information in the optic array (ϕ_k) to coordinate its behavior with respect to the layout of objects and surfaces in its environment *(G)*. Several studies demonstrate that acquiring this ability is dependent on specific visual experiences that define the coenetic variable (y_0) in this instance of medium-term adaptation. Precise definition of the coenetic variable depends on what component of visuomotor coordination is under consideration. For example, normally reared kittens, when held in the air, show a "visual placing response" (extension of the forelimbs) that is elicited by an approaching surface and, if the surface is interrupted by gaps, they will guide their forepaws to meet the solid parts of the surface. Hein and Held (1967) found, in kittens reared in a normal visual environment but denied sight of their limbs by a collar worn around the neck, that the elicited component develops normally but that the guided component is absent. The coenetic variable of adaptation apparently differs in the development of these two components of the placing response: For the elicited component, exposure to patterned light is sufficient for normal development (see Hein, Gower, & Diamond, 1970); for the guided component, visual experience of the limbs is also required (Hein & Held, 1967). This specificity between the coenetic variable and the behavioral skill whose development is being considered shows this example to be a relatively typical instance of learning, as previously characterized.

The visual placing response is an example of closed-loop reaching, in which the animal can see both its paw and the target. Under open-loop conditions, by contrast, the position of the target is indicated by a visual marker but the animal can see neither its own paw nor the target. Monkeys that are reared without sight of their hands show deficiencies in open-loop reaching that are similar to those shown by deprived kittens in the closed-loop task (Bauer & Held, 1975; Held & Bauer, 1967, 1974; Walk & Bond, 1971). Whereas normally reared animals reach directly for the target, guiding their reach with reference to the visual marker, deprived animals flail their arms wildly until they hit the target accidentally and then grasp it.

The open loop situation is interesting because the coenetic variable in the development of this skill (sight of the arms) is not the same as the environmental variable to which adaptation is later effected (specification of a hidden target position by visual information). This was also true, it will be recalled, in the case of species identification in ducklings, previously discussed in Section IV,A. A further point of interest is that the environmental variable ϕ is a complex relation between the pattern of optic texture and the position of graspable objects *with regard to the perceiver;* a full understanding of this instance of adaptation therefore awaits an animal-relevant style of description of the environment and of the optical structure that it generates, along the lines of that provided by Lee (1974) for the case of locomotion.

D. MEDIUM-TERM ADAPTATION AND ENVIRONMENTAL VARIABILITY

We have pointed out that learning permits adaptation to be effected to aspects of the environment that change rapidly in relation to evolutionary time-scales (see also Plotkin & Odling-Smee, 1979; Slobodkin, 1968;

Slobodkin & Rapoport, 1974). Identifying variation in ϕ is thus an appropriate concern for an ecological approach to the study of learning, since it indicates the possible selective value of particular learning skills and so helps to account for their evolution in the animals that possess them. The terms "indicates" and "possible" are used deliberately, for establishing the selective value of any phenotypic characteristic with any degree of certainty is fraught with difficulty, both conceptual and empirical (see Hinde, 1975; Lewontin, 1979).

In some of the cases we have discussed, it is possible to identify the environmental variability that might account for the adaptiveness of these learning abilities. Thus the relatively rapid speciation of waterfowl and songbirds might be offered as a reason for the involvement of learning in species identification and in song development, respectively (see Immelmann, 1975). In the case of stellar orientation in the indigo bunting, Emlen (1975b) has pointed out that the precession of the earth's axis of rotation, which changes the relationship between celestial and geographic directions at the rate of about 3° every 1000 years, might be identified as the evolutionary reason for the involvement of learning in the development of this skill.

We must, however, guard against the "adaptationist fallacy" of assuming that all characteristics of an organism *must* be explained by appeal to their possible adaptive benefits to their possessor (Gould & Lewontin, 1979; Lewontin, 1979). We have already pointed out that natural selection and evolutionary adaptation, while intimately related processes, are not identical. The ability of natural selection to produce precisely the adaptations that an animal requires and to make those adaptations optimally efficient is limited in a number of important ways (see Section IV,A). Consider, for instance, the limitation imposed by a lack of appropriate genetic variation in the population. A population that experiences a change in its environment, such as a new food source, may not possess the kind or amount of genetic variation required to evolve, say, a different adaptive tooth structure. It may, however, possess more genetic variation associated with the development of its feeding behavior, so that a learning skill can evolve enabling individuals to adapt to the characteristics of the food source on the basis of their own experience, *even though those characteristics may remain invariant over long periods of time*. In this case, the evolutionary reason for the possession of a learning skill has more to do with past genetic variation in the population than with the nature of environmental variability.

It is quite possible that different kinds and amounts of past genetic variation account for the different styles of song development in passerine birds (Marler & Mundinger, 1971); while some species show a strong dependence on exposure to a song model for normal song development, others develop almost completely normal adult song when raised in auditory isolation. No convincing adaptive explanation of these differences has yet been offered and it may be that none is needed or possible. They may simply reflect equivalent adaptive responses to similar selection pressures by genetically different populations (Lewontin, 1979).

This is not to say that attempts to provide adaptive explanations for the possession of learning skills are always misplaced. Where environmental variability is such that long-term adaptation cannot be effected, natural selection will tend to favor those individuals capable of effecting medium-term adaptive responses in the course of development and such learning skills may clearly be given an adaptive explanation. Providing such explanations, however, is far from straightforward and alternative, nonadaptive explanations should always be borne in mind. Perhaps the greatest danger of adaptive explanations is that while they are very easy to construct, they are usually very difficult or impossible to test. This increases the temptation to accept plausible adaptive stories in lieu of experimental demonstrations (Gould & Lewontin, 1979; Lewontin, 1979). Much more could be said on this important topic and these brief remarks are intended merely as a caveat for future work on this problem.

IV. Concluding Discussion

In this article we have presented a sketch of a metatheory, or conceptual framework, that can serve to guide an ecological approach to the study of learning. Our aim has been to set the problems of learning in a broader context, a context supplied largely by ecological and evolutionary theory, and to indicate the kinds of questions that might be asked about learning from the perspective thus provided. Foremost among the issues that emerge from this undertaking is an overarching concern with the relationship between the animal that learns and the environment that is learned about. This relationship is not an arbitrary one; each animal stands in a special, ecological relationship to a particular environment, one that it encounters as a result of its phylogeny and normal ontogeny. Defining the "natural environment" for particular species may be a difficult and demanding task, but the difficulty of this endeavor should not be allowed to overshadow its importance.

The theoretical significance that we attach to the relationships between animals and their (natural) environments is reflected in the account that we have given of the ecosystem, not the animal, as the minimal unit for the ecological study of learning. Within the ecosystem, animal and environment are defined in interdependent, coimplicative terms, as actor and econiche. We intend the unity of the ecosystem to be taken quite literally, not merely as an expression of good intent to be forsaken when attention is turned to the analysis of concrete examples of learning (see also Bentley, 1941), but this does not mean that we see the ecosystem as analytically impenetrable. Analytically, our attitude toward the ecosystem is somewhat akin to that of a modern neurophysiologist toward the brain: It is a unified system, with richly structured, nonarbitrary relationships among its many components but its complexity is such that experimental "dissection" is the only possible route to understanding its systemic properties. The implementation of this strategy toward the study of learning is clearly shown by the examples analyzed in Section V, in which experimental interference with the normal processes of development provides the key to understanding those processes.

A second issue that has concerned us deeply is that of the animal's adaptation to its environment and, in particular, of the unity of that adaptation. It is commonplace in biology to observe that the animal is an adapted whole, not an assemblage of adapted elements, and we would extend this observation to include the fact of temporal as well as spatial unity: An animal is not a succession of instants; it is an extended *event* and in striving for a unified account of adaptation on different time-scales, we have given explicit recognition to this fact.

As a conceptual tool for achieving our goal of a unified account of adaptation we have employed the formal model of adaptation developed by Sommerhoff (1950, 1969), a model that has suffered undeserved obscurity during the 30 years since its first publication. We have employed the model as an analytic rather than a predictive tool, analysis being the aim of this article as well as Sommerhoff's original intent. The advantages of employing this formalism are, we hope, evident. It allows us to define, in precise and unambiguous terms, the elements of any instance of adaptation and the relationships between those elements. Most importantly, it provides a vocabulary in which to phrase questions about adaptation, and about learning in particular, that are both ecologically motivated and amenable to experimental investigation. As noted in Section I, the primary aim of metatheoretical inquiry is to raise particular kinds of questions, in our case ecological questions about learning, and Sommerhoff's model provides an admirable tool in that regard.

As far as we have been able to define them in the space of this article, the questions and issues that are raised by the ecological approach to learning are importantly different from those raised by more traditional approaches. Questions concerning the relationships of learning to other forms of adaptation, including those on both similar and different time-scales, are given a prominence by the unified view of adaptation that is largely lacking in traditional approaches. The problem of description of the environment, as an empirical rather than a logical problem, is not one that has traditionally been raised but it assumes major significance in the ecological approach.

Traditional approaches to the study of learning emphasize, often to the virtual exclusion of other considerations, analysis of the mechanisms of learning. While we have paid rather little attention to questions of mechanism (an emphasis in line with our metatheoretical aims) it is clear that this is an important area for future investigation. It is important to recognize, however, that since we have proposed a view of learning as a *direct* adaptive response, the kinds of mechanism that might be proposed to account for it will probably look very different from the *indirect* mechanisms (involving association, memory, propositional structures and the like) that are currently popular. Gottlieb's (1976) three roles of experience represent an important step toward understanding the mechanisms of direct learning and further work on this problem is urgently needed.

The difference between the two approaches may also be seen in the traditionally important questions that are not raised by the ecological approach. Most of the important issues in current psychology of learning presuppose an associationist account of learning (see Jenkins, 1979). The ecological approach does not involve associationism and so issues such as the role of reinforcement, the nature of Pavlovian-operant interactions, and stimulus-response specificity simply do not arise. Some traditionally important issues (of which the learning-performance distinction may be one) do seem likely to be raised as an ecological account of learning develops, but we doubt that they will take the same form as they do in association theory.

We make no claim to have considered in this article all of the issues that are involved in an ecological approach to the study of learning, nor to have exhausted the problems inherent in those issues that we have considered. Our main hope is that by taking a broad perspective view of the nature of learning we have at least been able to cast some important issues in sharp relief, and so point in the direction of their resolution.

REFERENCES

Balderrama, N., & Maldonado, H. Ontogeny of the behavior in the praying mantis. *Journal of Insect Physiology,* 1973, **19,** 319-336.

Baldwin, J. M. *Development and evolution.* London: Macmillian, 1902.

Bateson, P. P. G. Specificity and the origins of behavior. In J. S. Rosenblatt, R. A. Hinde, C. G. Beer, & M. C. Busnel (Eds.), *Advances in the study of behavior* (Vol. 6). New York: Academic Press, 1976.

Bauer, J. A., & Held, R. Comparison of visually guided reaching behavior in normal and deprived infant monkeys. *Journal of Experimental Psychology: Animal Behavior Processes,* 1975, **1,** 298-308.

Beck, S. D. Environmental photoperiod and the programming of insect development. In E. T. Drake (Ed.), *Evolution and environment.* New Haven: Yale University Press, 1968.

Bell, P. R. The movement of plants in response to light. In P. R. Bell (Ed.), *Darwin's biological work: Some aspects reconsidered.* New York: Wiley, 1959.

Bentley, A. F. The factual space and time of behavior. *Journal of Philosophy,* 1941, **38,** 477-485.

Bock, W. J., & von Wahlert, G. Adaptation and the form-function complex. *Evolution,* 1965, **19,** 269-299.

Bolles, R. C. Species-specific defense reactions and avoidance learning. *Psychological Review,* 1970, **77,** 32-48.

Bolles, R. C. Species-specific defense reactions. In F. R. Brush (Ed.), *Aversive conditioning and learning.* New York: Academic Press, 1971.

Bolles, R. C. The avoidance learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 6). New York: Academic Press, 1972.

Collias, N. E., & Collias, E. C. Some mechanisms of family integration in ducks. *Auk,* 1956, **73,** 378-400.

Cornman, J. W. *Perception, common sense, and science.* New Haven: Yale University Press, 1975.

Cummins, R. Programs in the explanation of behavior. *Philosophy of Science,* 1977, **44,** 269-287.

Darwin, C. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London: Murray, 1859.

DeBeer, G. *Embryos and ancestors* (3rd. ed.). London: Oxford University Press, 1958.

Dewey, J., & Bentley, A. F. *Knowing and the known.* Boston: Beacon, 1949.

Dobzhansky, Th. Biological adaptation. *Scientific Monthly,* 1942, **55,** 391-402.

Dobzhansky, Th. What is an adaptive trait? *American Naturalist,* 1956, **90,** 337-347.

Dobzhansky, Th. On some fundamental concepts of Darwinian biology. In Th. Dobzhansky, M. K. Hecht, & W. M. Steere (Eds.), *Evolutionary biology* (Vol. 2). New York: Appleton, 1968.

Elton, C. *Animal ecology.* London: Sidgwick & Jackson, 1927.

Emlen, J. M. *Ecology: An evolutionary approach.* Reading, Mass.: Addison-Wesley, 1973.

Emlen, S. T. Migratory orientation in the indigo bunting, *Passerina cyanea.* Part I: Evidence for the use of celestial cues. *Auk,* 1967, **84,** 309-342. (a)

Emlen, S. T. Migratory orientation in the indigo bunting, *Passerina cyanea.* Part II: Mechanism of celestial orientation. *Auk,* 1967, **84,** 463-489. (b)

Emlen, S. T. The development of migratory orientation in young indigo buntings. *Living Bird,* 1969, **8,** 113-126.

Emlen, S. T. Celestial rotation: Its importance in the development of migratory orientation. *Science,* 1970, **170,** 1198-1201.

Emlen, S. T. The ontogenetic development of orientation capabilities. In S. R. Galler, K. Schmidt-Koenig, G. J. Jacobs, & R. F. Belleville (Eds.), *Animal orientation and navigation.* Washington, D. C.: NASA (SP-262), 1972.

Emlen, S. T. Migration: Orientation and navigation. In D. S. Farner & J. R. King (Eds.), *Avian biology* (Vol. 5). New York: Academic Press, 1975. (a)

Emlen, S. T. The stellar-orientation system of a migratory bird. *Scientific American,* 1975, **233**(2), 102-111. (b)

Evans, F. C. Ecosystem as the basic unit in ecology. *Science,* 1956, **123,** 1127-1128.

Fitch, H., & Turvey, M. T. On the control of activity: Some remarks from an ecological point of view. In R. W. Christiana (Ed.), *Psychology of motor behavior and sports*. Urbana, Ill.: Human Kinetics Press, 1979.

Fodor, J. A. *Psychological explanation: An introduction to the philosophy of psychology*. New York: Random House, 1968.

Fodor, J. A. Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, 1980, in press.

Fowler, C. A. *Timing in relation to speech perception*. Ph.D. dissertation, University of Connecticut, 1977.

Frazzetta, T. H. *Complex adaptations in evolving populations*. Sunderland, Mass.: Sinauer. 1975.

Ghiselin, M. T. On semantic pitfalls of biological adaptation. *Philosophy of Science*, 1966, **33**, 147-153.

Ghiselin, M. T. *The economy of nature and the evolution of sex*. Berkeley: University of California Press, 1974.

Gibson, J. J. *The perception of the visual world*. Boston: Houghton, 1950.

Gibson, J. J. Visually guided locomotion and visual orientation in animals. *British Journal of Psychology*, 1958, **49**, 182-194.

Gibson, J. J. *The senses considered as perceptual systems*. Boston: Houghton, 1966.

Gibson, J. J. The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Towards an ecological psychology*. Hillsdale, N.J.: Erlbaum, 1977.

Gibson, J. J. *The ecological approach to visual perception*. Boston: Houghton, 1979.

Gottlieb, G. Imprinting in relation to parental and species identification by avian neonates. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 345-356.

Gottlieb, G. *Development of species identification in birds: An inquiry into the prenatal determinants of perception*. Chicago: University of Chicago Press, 1971.

Gottlieb, G. Development of species identification in ducklings: I. Nature of perceptual deficit caused by embryonic auditory deprivation. *Journal of Comparative and Physiological Psychology*, 1975, **89**, 387-399. (a)

Gottlieb, G. Development of species identification in ducklings: II. Experiential prevention of perceptual deficit caused by embryonic auditory deprivation. *Journal of Comparative and Physiological Psychology*, 1975, **89**, 675-684. (b)

Gottlieb, G. Development of species identification in ducklings: III. Maturational rectification of perceptual deficit caused by auditory deprivation. *Journal of Comparative and Physiological Psychology*, 1975, **89**, 899-912. (c)

Gottlieb, G. The roles of experience in the development of behavior and the nervous system. In G. Gottlieb (Ed.), *Studies on the development of behavior and the nervous system* (Vol. 3: *Neural and behavioral specificity*). New York: Academic Press, 1976.

Gottlieb, G. Development of species identification in ducklings: IV. Change in species-specific perception caused by auditory deprivation. *Journal of Comparative and Physiological Psychology*, 1978, **92**, 375-387.

Gottlieb, G. Development of species identification in ducklings: V. Perceptual differentiation in the embryo. *Journal of Comparative and Physiological Psychology*, 1979, **93**, 831-854.

Gottlieb, G. Development of species identification in ducklings: VI. Specific embryonic experience required to maintain species-typical perception in Peking ducklings. *Journal of Comparative and Physiological Psychology*, in press.

Gottlieb, G., & Vandenbergh, J. G. Ontogeny of vocalization in duck and chick embryos. *Journal of Experimental Zoology*, 1968, **168**, 307-325.

Gould, S. J. *Ontogeny and phylogeny*. Cambridge: Harvard University Press, 1977.

Gould, S. J., & Lewontin, R. C. The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist program. *Proceedings of the Royal Society of London, Series B*, 1979, **205**, 581-598.

Grinnell, J. Geography and evolution. *Ecology*, 1924, **5**, 225-229.

Hamilton, W. D. The genetical evolution of social behavior, I & II. *Journal of Theoretical Biology,* 1964, **7,** 1-52.

Haugeland, J. The nature and plausibility of cognitivism. *Behavioral and Brain Sciences,* 1978, **2,** 215-260 [with commentary].

Hein, A., Gower, E. C., & Diamond, R. M. Exposure requirements for developing the triggered component of the visual placing response. *Journal of Comparative and Physiological Psychology,* 1970, **73,** 188-192.

Hein, A., & Held, R. Dissociation of the visual placing response into elicited and guided components. *Science,* 1967, **158,** 390-392.

Held, R., & Bauer, J. A. Visually guided reaching in infant monkeys after restricted rearing. *Science,* 1967, **155,** 718-720.

Held, R., & Bauer, J. A. Development of sensorially-guided reaching in infant monkeys. *Brain Research,* 1974, **71,** 247-258.

Henderson, L. J. *The fitness of the environment.* New York: Macmillan, 1913.

Hinde, R. A. The concept of function. In G. Baerends, C. G. Beer, & A. Manning (Eds.), *Function and evolution in behaviour: Essays in honour of Professor Niko Tinbergen, FRS.* London: Oxford University Press (Clarendon), 1975.

Ho, M. W., & Saunders, P. T. Beyond neo-Darwinism—an epigenetic approach to evolution. *Journal of Theoretical Biology,* 1979, **78,** 573-591.

Holling, C. S. The analysis of complex population processes. *Canadian Entomologist,* 1964, **96,** 335-347.

Humphrey, G. *The nature of learning in its relation to the living system.* London: Kegan, 1933.

Hutchinson, G. E. Concluding remarks. *Cold Spring Harbor Symposia in Quantitative Biology,* 1957, **22,** 415-427.

Hutchinson, G. E. *The ecological theater and the evolutionary play.* New Haven: Yale University Press, 1967.

Immelmann, K. Ecological significance of imprinting and early learning. *Annual Review of Ecology and Systematics,* 1975, **6,** 15-37.

Jenkins, H. M. Animal learning and behavior theory. In E. Hearst (Ed.), *The first century of experimental psychology.* Hillsdale, N. J. Erlbaum, 1979.

Johansson, G. Projective transformations as determining visual space perception. In R. B. McLeod & H. L. Pick (Eds.), *Perception: Essays in honor of J. J. Gibson.* Ithaca, N. Y.: Cornell University Press, 1974.

Johnston, T. D. *Ecological dimensions of learning: A study of motor skill acquisition in infant primates.* Ph.D. dissertation, University of Connecticut, 1978.

Johnston, T. D. *Contrasting approaches to a theory of learning.* Submitted.

Kalat, J. W. Biological significance of food aversion learning. In N. W. Milgram, L. Krames, & T. M. Alloway (Eds.), *Food aversion learning.* New York: Plenum, 1977.

Kamin, L. J. "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation.* Miami: University of Miami Press, 1968.

Kamin, L. J. Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior.* New York: Appleton, 1969.

Konishi, M., & Nottebohm, F. Experimental studies in the ontogeny of avian vocalizations. In R. A. Hinde (Ed.), *Bird vocalizations.* London: Cambridge University Press, 1969.

Kuhn, T. *The structure of scientific revolutions.* Chicago: University of Chicago Press, 1962.

Kuo, Z.-Y. *The dynamics of behavior development.* New York: Random House, 1967.

Lakatos, I. Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge.* London: Cambridge University Press, 1970.

Leathwood, P. Influence of early undernutrition on behavioral development and learning in rodents. In G. Gottlieb (Ed.), *Studies on the development of behavior and the nervous system* (Vol. 4: *Early influences*). New York: Academic Press, 1978.

Lee, D. N. Visual information during locomotion. In R. B. McLeod & H. L. Pick (Eds.), *Perception: Essays in honor of J. J. Gibson.* Ithaca, N. Y.: Cornell University Press, 1974.

Lee, D. N. A theory of visual control of braking based on information about time to collision. *Perception,* 1976, **5,** 437–459.

Lee, D. N., & Aronson, E. Visual proprioceptive control of standing in human infants. *Perception and Psychophysics,* 1974, **15,** 529–532.

Lee, D. N., & Lishman, R. Visual control of locomotion. *Scandanavian Journal of Psychology,* 1977, **18,** 224–230.

Lehrman, D. S. A critique of Konrad Lorenz's theory of instinct. *Quarterly Review of Biology,* 1953, **28,** 337–363.

Lewontin, R. C. The adaptations of populations to varying environments. *Cold Spring Harbor Symposia in Quantitative Biology,* 1957, **22,** 395–408.

Lewontin, R. C. *The genetic basis of evolutionary change.* New York: Columbia University Press, 1974.

Lewontin, R. C. Adaptation. *Scientific American,* 1978, **239**(3), 212–230.

Lewontin, R. C. Sociobiology as an adaptationist program. *Behavioral Science,* 1979, **24,** 5–14.

Løvtrup, S. *Epigenetics: A treatise on theoretical biology.* New York: Wiley, 1974.

Mackintosh, N. S., & Sutherland, N. J. *Mechanisms of animal discrimination learning.* New York: Academic Press, 1971.

Maldonado, H., Levin, L., & Barros Pita, J. C. Hit distance and the predatory strike of the praying mantis. *Zeitschrift für vergleichende Physiologie,* 1967, **56,** 236–257.

Maldonado, H., Rodriguez, E, & Balderrama, N. How mantids gain insight into the new maximum catching distance after each ecdysis. *Journal of Insect Physiology,* 1974, **20,** 591–603.

Marler, P. R., & Mundinger, P. Vocal learning in birds. In H. Moltz (Ed.), *The ontogeny of vertebrate behavior.* New York: Academic Press, 1971.

Mason, H. L., & Langenheim, J. H. Language analysis and the concept *environment. Ecology,* 1957, **38,** 325–340.

Medawar, P. B. Problems of adaptation. *New Biology,* 1951, **11,** 10–26.

Mettler, L. E., & Gregg, T. G. *Population genetics and evolution.* New York: Prentice-Hall, 1969.

Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, 1960.

Mittelstaedt, H. Prey capture in mantids. In B. T. Scheer (Eds.), *Recent advances in invertebrate physiology.* Eugene: University of Oregon Press, 1957.

Mundle, C. W. K. *Perception: Facts and theories.* London: Oxford University Press, 1971.

Nakayama, K., & Loomis, J. M. Optical velocity patterns, velocity sensitive neurons and space perception: A hypothesis. *Perception,* 1974, **3,** 63–80.

Nottebohm, F. Ontogeny of bird song. *Science,* 1970, **167,** 950–956.

Odum, E. P. *Fundamentals of ecology.* Philadelphia: Saunders, 1953.

Odum, E. P. *Fundamentals of ecology* (2nd. ed.). Philadelphia: Saunders, 1959.

Plotkin, H. C., & Odling-Smee, F. J. Learning, change, and evolution: An enquiry into the teleonomy of learning. In J. S. Rosenblatt, R. A. Hinde, C. G. Beer, & M. C. Busnel (Eds.), *Advances in the study of behavior* (Vol. 10). New York: Academic Press, 1979.

Purton, A. C. Ethological categories of behaviour and some consequences of their conflation. *Animal Behaviour,* 1978, **26,** 653–670.

Rescorla, R. A. Informational variables in Pavlovian conditioning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 6). New York: Academic Press, 1972.

Rescorla, R. A., & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. R. Black & W. F. Prokasy (Eds.), *Classical conditioning II*. New York: Appleton, 1972.

Rosenzweig, M. R., & Bennett, E. L. Experiential influences on brain anatomy and brain chemistry in rodents. In G. Gottlieb (Ed.), *Studies on the development of behavior and the nervous system* (Vol. 4: *Early influences*). New York: Academic Press, 1978.

Rozin, P., & Kalat, J. W. Specific hungers and poison avoidance as adaptive specializations of learning. *Psychological Review*, 1971, **78**, 459–486.

Ryle, G. *The concept of mind*. London: Hutchinson, 1949.

Schiff, W. Perception of impending collision: A study of visually directed avoidant behavior. *Psychological Monographs*, 1965, **79** (Whole No. 604).

Schmalhausen, I. *Factors of evolution*. Philadelphia: Blakiston, 1949.

Schneirla, T. C. Interrelationships of the "innate" and the "acquired" in instinctive behavior. In P.-P. Grasse (Ed.), *L'Instinct dans le comportement des animaux et de l'homme*. Paris: Masson, 1956.

Schneirla, T. C. Aspects of stimulation and organization in approach/withdrawal processes underlying vertebrate social development. In D. S. Lehrman, R. Hinde, & E. Shaw (Eds.), *Advances in the study of behavior* (Vol. 1). New York: Academic Press, 1965.

Seligman, M. E. P. On the generality of the laws of learning. *Psychological Review*, 1970, **77**, 406–418.

Shaw, R. E., & Turvey, M. T. Coalitions as models for ecosystems: A realist perspective on perceptual organization. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, N. J.: Erlbaum, 1981, in press.

Shaw, R. E., Turvey, M. T., & Mace, W. M. Ecological psychology: The consequence of a commitment to realism. In W. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes, II*. Hillsdale, N. J.: Erlbaum, 1981, in press.

Shettleworth, S. J. Constraints on learning. In D. S. Lehrman, R. A. Hinde, & E. Shaw (Eds.), *Advances in the study of behavior* (Vol. 4). New York: Academic Press, 1972.

Slobodkin, L. B. Toward a predictive theory of evolution. In R. C. Lewontin (Ed.), *Population biology and evolution*. Syracuse, N. Y.: Syracuse University Press, 1968.

Slobodkin, L. B., & Rapoport, A. An optimum strategy of evolution. *Quarterly Review of Biology*, 1974, **49**, 181–200.

Sommerhoff, G. *Analytical biology*. London: Oxford University Press, 1950.

Sommerhoff, G. The abstract characteristics of living systems. In F. E. Emery (Ed.), *Systems thinking*. Baltimore, Md.: Penguin, 1969.

Stearns, S. C. The evolution of life history traits: A critique of the theory and a review of the data. *Annual Review of Ecology and Systematics*, 1977, **8**, 145–171.

Stern, J. T. The meaning of "adaptation" and its relation to the phenomenon of natural selection. In Th. Dobzhansky, I. M. Hecht, & W. C. Steere (Eds.), *Evolutionary biology* (Vol. 4). New York: Appleton, 1970.

Strong, D. R., & Ray, T. S. Host tree location behavior of a tropical vine *(Monstera gigantea)* by skototropism. *Science*, 1975, **190**, 804–806.

Tolman, E. C. There is more than one kind of learning. *Psychological Review*, 1949, **56**, 144–155.

Turvey, M. T. Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Towards an ecological psychology*. Hillsdale, N. J.: Erlbaum, 1977. (a)

Turvey, M. T. Contrasting orientations to the theory of visual information processing. *Psychological Review*, 1977, **84**, 67–88. (b)

Turvey, M. T., & Shaw, R. E. The primacy of perceiving: An ecological reformulation of perception for understanding memory. In L. G. Nilsson (Ed.), *Prospectus on memory research: Essays* *in honor of Uppsala University's 500th anniversary.* Hillsdale, N. J.: Erlbaum, 1979.

Turvey, M. T., Shaw, R., & Mace, W. M. Issues in the theory of action: Degrees of freedom, coordinative structures, and coalitions. In J. Requin (Ed.), *Attention and performance VII.* Hillsdale, N. J.: Erlbaum, 1978.

Vandemeer, J. H. Niche theory. *Annual Review of Ecology and Systematics,* 1972, **3,** 107–132.

Waddington, C. H. *The strategy of the genes.* London: Allen & Unwin, 1957.

Walk, R. D., & Bond, E. K. The development of visually guided reaching in monkeys reared without sight of the limbs. *Psychonomic Science,* 1971, **23,** 115–116.

Warren, R. The perception of egomotion. *Journal of Experimental Psychology: Human Perception and Performance,* 1976, **2,** 448–456.

Whittaker, R. H., Levin, S. A., & Root, R. B. Niche, habitat, and ecotope. *American Naturalist,* 1973, **107,** 321–338.

Williams, G. C. *Adaptation and natural selection: A critique of some current evolutionary thought.* Princeton, N. J.: Princeton University Press, 1966.

This Page Intentionally Left Blank

SAM: A THEORY OF PROBABILISTIC SEARCH OF ASSOCIATIVE MEMORY

Jeroen G. W. Raaijmakers

UNIVERSITY OF NIJMEGEN, NIJMEGEN, THE NETHERLANDS

Richard M. Shiffrin

INDIANA UNIVERSITY, BLOOMINGTON, INDIANA

We shall introduce in this article a theory of retrieval from long-term memory, and present a number of applications to data from paradigms involving free recall, categorized free recall, and paired-associate recall. The theory combines elements of probabilistic search theory (e.g., Shiffrin, 1970) and associative network theory (e.g., Anderson, 1972). It posits cue-dependent probabilistic search of an associative long-term

ISBN 0-12-543314-X

memory network, and is denoted SAM, for Search of Associative Memory.

Our goals, given the length of this article, are limited in scope. The general theory is surveyed briefly, but the reader is referred to Raaijmakers and Shiffrin (1981) for a detailed discussion of the underlying basis of the theory. Then a quantitative simulation model of SAM will be described. This model is used, in essentially intact form, in many cases with no changes in parameters, to fit data from a variety of memory paradigms. To reduce the article's length these paradigms will be restricted to free and cued recall tasks for lists of singly presented items to be remembered.

I. A Search Theory for Retrieval from Associative Memory

A. THE STRUCTURE OF LONG-TERM STORE (LTS)

Long-term store (LTS) is held to be a richly interconnected network, with numerous levels, stratifications, categories, and trees, containing varieties of relationships, schemata, frames, and associations. Roughly speaking, all elements of memory are connected to all others, directly or indirectly (though perhaps quite weakly).

The "objects" of memory are defined by the task and the level of analysis pursued by the investigator. The boundary of a memory object is seldom clearly defined. For example, a "word," a "letter," and a "story" may be memory objects in different tasks; each consists of a complex bundle of informational elements, associations, and relations. In the present article the level of analysis is chosen so that the "word image," or some other similarly complex and distinct entity (such as a picture), is the basic object of memory. Even though a memory object has no clear boundaries, it can make sense to distinguish such objects from each other, in the sense that interconnections between elements and features will be stronger and more numerous within one object than between objects. Thus a memory object tends to be a relatively unitized entity.

We propose that memory structure at a given level of analysis be summarized in a retrieval structure. This structure contains retrieval strengths between the possible probe cues and the objects in memory. These strengths represent an average associative relationship between probe cues and memory objects, ignoring details such as the kind of relationships involved. As we shall indicate below, these strengths are used in a simple ratio rule to determine the probability that a particular object will be elicited from memory when long-term memory is probed with a given set of cues. The retrieval structure is designed to capture those aspects of the memory structure that are important for retrieval. The only restriction on these strengths is that they be positive numbers. Such a structure is a rich enough representation for our retrieval model to predict many results from a variety of paradigms.

In the tasks treated in this article, the memory objects will typically be combinations of word features and contextual features, called "images." The important role played by temporal and contextual information is understandable in light of the tasks, requiring memory that a word was presented during a particular list. It would not do to let the memory image consist of word information without temporal context, since the strong preexperimental strengths between such images would mask the relatively small increments in strength that would occur due to presentation in a single list. Temporal-context, separate from word information, may be used as a cue to probe such a memory structure, or combinations of context with words may be used as cue sets to probe memory. Presumably, context alone is used as a cue when no words are available, as might be the case at the start of free recall.

Although SAM does not require a particular memory representation, it is useful to give one simplified representation to illustrate our main points. Figure 1 schematizes associations that might be formed after study of a five-item list. The item information, and the context information associated to the item information, are enclosed by solid lines. The strength of association of context to an image, when context is used as a cue, is given by the solid arrows. The solid arrows point to the item information since in many tasks the "name" of the sampled image is required. The dashed regions enclosing both context and item features indicate images of an item within the present context, i.e., the memory objects. These images are associated to other such images and the strength and direction of these interitem associations are indicated by the dashed arrows. (Associations between features are complex and are not shown in this figure; also not shown are residual associations between items not rehearsed together.)

In general, we prefer to treat multiple representations of the "same" item as separate images, each with its particular temporal-contextual elements. However, these images may be closely associated due to their pool of shared features, so that a type-token (e.g., Anderson & Bower, 1973) or episodic-semantic (e.g., Tulving, 1972) distinction may still be maintained. The set of common semantic features in many images can be considered the "type" or semantic image.

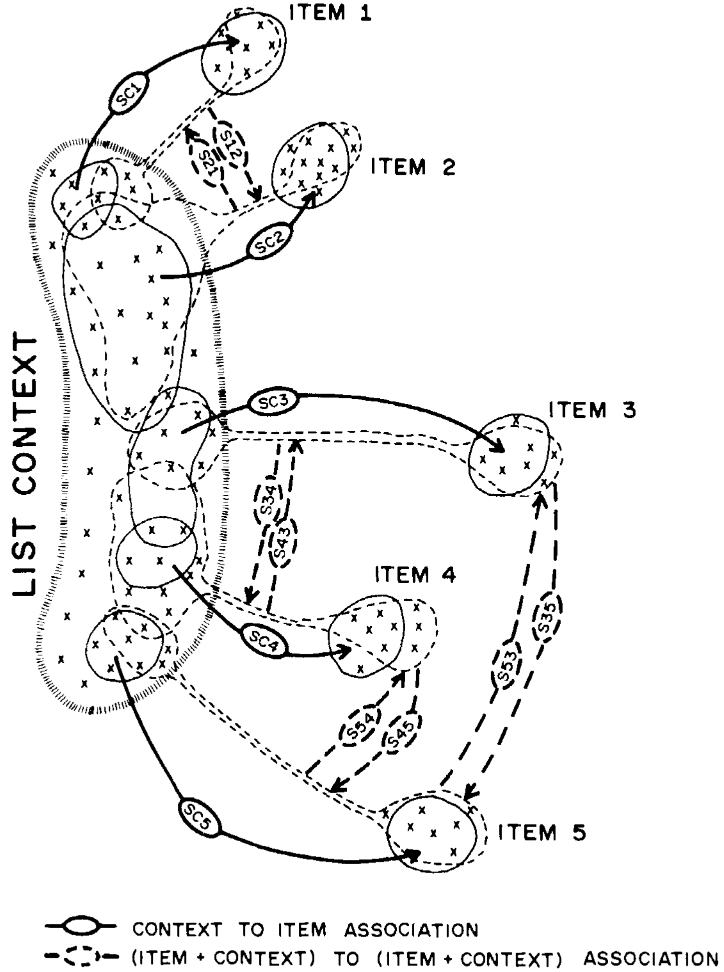


Fig. 1. A schematic depiction of the associative network in long-term store after study of a five-item list. Shown are the directional associations (dashed arrows) between images (each consisting of item plus context); also shown are the associations between context elements when used as cues, and the various images (solid arrows). Residual associations between images of items not rehearsed together are not shown. Xs refer to various features.

The final point to be emphasized is the permanence of long-term memory. We assume that information may be entered in, and added to, the long-term store, but not deleted from, or subtracted from the store. Forgetting is thus assumed to occur as a result of retrieval failure; factors governing retrieval failure will be discussed later.

B. RETRIEVAL FROM LONG-TERM STORE

The most important feature of the retrieval system is cue-dependence (see Tulving, 1974). Probe cues, whether consciously selected or not, govern each stage of the memory search. The degree to which an image in memory is associated to the set of probe cues, in comparison with the degrees to which other images are associated to the set of probe cues, determines the probability that that item will be selected at that moment in the memory search.

It is assumed that the retrieval system is noisy and inherently probabilistic; for a given memory structure and set of probe cues, the image selected from memory is a random variable. It is easy to misinterpret such a statement and ascribe more randomness to the retrieval system than is, in fact, present. The strength may be such that one image is far more likely to be selected than any other. Furthermore, the subject can control the search by changing the probe cues as needed. Nevertheless, the inherently random nature of the search has important consequences; for example, images that are sampled at one point in the search may be resampled later, especially if the probe cues are not changed.

The retrieval system as a whole is an extension of that proposed by Shiffrin (1970). It envisions retrieval as a memory search proceeding in a series of discrete steps, each step involving a selection, or sample, of an image from long-term store. The substages within any one step are depicted in Fig. 2. Retrieval begins with some question the subject needs to answer regarding the contents of long-term store. This may be as simple as "what is another word on the list most recently presented?" In the most general case, a retrieval plan will next be generated to guide the search for the answer. Initially, the plan may be somewhat vague by intention, in the hope that later phases of the search will be guided by information located in earlier phases. The plan includes such things as an initial decision whether to search long-term store, how to search (for instance, in a temporal order, or by an alphabetic strategy), how to choose probe cues (for instance, should recalled information be used as probe cues?), what combinations of probe cues should be employed, with what weights, whether to employ the same probe cues on successive loops of the search or whether to alter the cues, whether to search first for preliminary cues to guide later search, and how long to search (i.e., how many loops of the search process are expected). Of course, the plan itself is constructed on the basis of the information in the test query, the information currently available in short-term memory, and information retrieved from long-term memory; the long-term information may be concerned

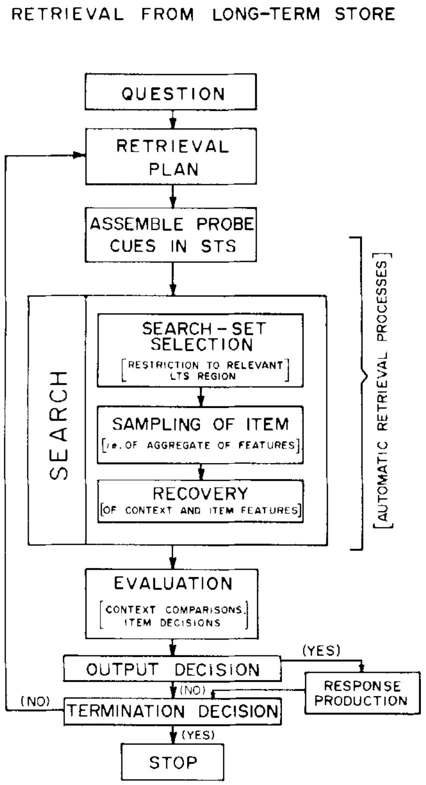


Fig. 2. A generalized depiction of the various phases of retrieval in the theory.

with search plans, previous successful plans in similar situations, and so forth (see Williams, 1977, for a discussion of retrieval plans).

Next, on the basis of the retrieval plan, the subject assembles probe cues to be used in retrieval. Generally, these cues will include: (1) information the subject has about the context at the time of study, (2) context representative of the moment of test (although these cues may not be useful or desired), (3) information from the test question, (4) information retrieved earlier in the search, and (5) information generated during construction of the retrieval plan. It is almost certainly the case, however, that there will be limitations on the amount of information that may be combined effectively into a set of probe cues. Perhaps the number of probe cues that may be used has an upper limit, or perhaps the various cues are weighted in importance (and effect), the sum of the weights being limited.

We argue that long-term memory images and probe cues are quite distinct. For example, suppose "horse" has been placed in long-term memory at the time of study, and then a recognition test is given with "horse" as the test item. The image will consist of "horse-at-study plus study context" and the test cues will consist of "horse-at-test" along with "test context" (the encodings of horse may differ in the two instances). Thus if the memory image is sampled, it may be evaluated alone or compared with the probe cues. However, these two entities will usually be strongly associated due to their large pool of common information; it is this fact that makes it likely that the cue "horse" will cause the image "horse" to be sampled. We shall not deal with recognition in this article, but even in recall tasks note that the image corresponding to a word cue may often be sampled (though such a sample will not be useful).

The next phases of the retrieval process concern sampling and recovery. As opposed to the other stages, these stages are largely automatic and not under direct control of the subject. They determine what image is sampled and how much of the information in (or perhaps near) the sampled image becomes available to the subject for evaluation and decision making.

An image has a probability of being sampled that is determined by the associative strength relating the set of probe cues to the image, in comparison with the strengths relating all other images to the set of probe cues. (This rule will be quantified shortly.) In fact, almost all images in long-term memory will have such low strengths of association to the cues, that their sampling probabilities will be vanishingly small. The relatively small set of images with nonnegligible sampling probabilities is denoted the "search-set." It is therefore convenient (especially when incorporating the model in a computer simulation) to separate the sampling phase into two parts: first, a restriction to the search-set; second, an appropriate probabilistic choice from the search-set. The choice of search-set is generally determined by task considerations. For example, if a subject is asked to recall a just-presented list, the search-set might be assumed to consist of the images of all the words in that list (or perhaps of all the words in the session, if it is necessary to predict intrusions).

When an image is sampled, its features will tend to become activated. It is assumed that the stronger is the association between the selected image and the probe cues, the larger will be the proportion of image elements that will be activated and made available to the subject's evaluation and decision-making mechanisms. This process is termed "recovery." It may well be that the particular elements recovered from a given image for a given set of cues are fixed (at least for the short-run), so that the same elements will be recovered if the same image is sampled several times in succession.

Once a given set of informational elements has been recovered, the subject carries out evaluations and makes appropriate decisions. Such evaluations include deciding what is the verbal "name" of the sampled image, whether the sampled image was indeed on the list being tested, whether the sampled image matches the test cue (in a recognition test), etc. The subject also decides whether he has succeeded in his search, whether a response should be output, and whether the search should be continued. If the search is continued, the process loops back to the retrieval plan to start the next step in the retrieval process.

C. QUANTITATIVE SAMPLING AND RECOVERY RULES

Let us begin by positing an $N + 1$ by N matrix, with every possible memory image in the search set (N of them) given horizontally, and every possible individual cue ($N + 1$ of them, including context) given vertically. The cues, excepting the context cue, correspond to the stored images in a one-to-one fashion. Thus each image represents an item that could be used as a cue. Let the matrix (Fig. 3) contain a strength (of association) between each cue and each image. Let $S_T(Q_i, I_j)$ denote this strength between Q_i and image I_j. (The T indicates that these are the strengths that apply at Test.) Call this a "retrieval structure."

The sampling assumption may now be stated as follows:

$$P_S(I_i | Q_1, Q_2 \ldots Q_M) = \frac{\prod_{j=1}^{M} \{S_T(Q_j, I_i)\}^{W_j}}{\sum_{k=1}^{N} \prod_{j=1}^{M} \{S_T(Q_j, I_k)\}^{W_j}} \tag{1}$$

The term on the left indicates the probability of sampling image I_i given cues $Q_1, \ldots, Q_M$ are used in combination as a probe set. The W_j in the right-hand expression are weights assigned to the different cues representing their relative saliency, or importance (or overlap, or similarity). (In the applications in this article it was unnecessary to assume unequal weights, and the W_j were all set to 1.0.)

The key to the present approach is the method used to combine cues: the strengths to the different cues are multiplied (perhaps in weighted fashion), and the ratio rule (Luce, 1959) applied to the products. This multiplicative feature has the useful and important consequence that it allows focusing of the search. The images with the highest probability of being sampled are those with the highest product of strengths, and hence those that tend to be strongly associated to *all* of the cues. The sampled image tends to come from the intersection of the sets of images strongly

IMAGES

CUES

$$\begin{array}{c|cccc|}
 & I_1 & I_2 & \cdots & I_N \\ \hline
Q_{i_0} & S_T(Q_{i_0},I_1) & S_T(Q_{i_0},I_2) & & S_T(Q_{i_0},I_N) \\
Q_{i_1} & S_T(Q_{i_1},I_1) & S_T(Q_{i_1},I_2) & & S_T(Q_{i_1},I_N) \\
\vdots & & & & \\
Q_{i_N} & S_T(Q_{i_N},I_1) & S_T(Q_{i_N},I_2) & & S_T(Q_{i_N},I_N) \\ \hline
\end{array}$$

Fig. 3. The "test matrix": the matrix of strengths which determine the probabilities of selection and recovery of list images (horizontal margin) when different cues (vertical margin) are used in the cue set. Entries in the cells are strengths from individual cues to individual images; when multiple cues are used in the cue-set, then the strengths are combined according to Eqs. (1) and (2) in the text. Q_{i_0} refers to the context cue; Q_{i_m} refers to the word whose image is I_m.

associated to each cue separately. By contrast, an additive combination rule could give a high probability of sampling an image if only one cue strength is high (even if all the others are zero).

Consider next the recovery process. We give here the recovery rule when the subject's task is to generate the "name" of the word encoded in the selected image:

$$P_R(I_i|Q_1,Q_2 \ldots Q_M) = 1 - \exp\{-\sum_{j=1}^{M} W_j S_T(Q_j, I_i)\} \tag{2}$$

The expression on the left represents the probability of recovering enough information to correctly give the "name" encoded in image I_i, which has just been selected using probe cues Q_1 to Q_M. The right-hand expression is somewhat arbitrary mathematically, though it does capture a number of features we consider desirable for a recovery rule in this case. First, the stronger the strength to any one cue and the stronger the summed strengths to all cues, the more likely is recovery. Second, the larger is a cue weight, the more the strength to that cue will affect recovery. Note that recovery obeys an additive rather than multiplicative rule, so that recovery probability will be high if even one weighted strength is high. Third, the probabilities will range from 0 to 1 as the sum of the strengths ranges from 0 to ∞. Note that both these sampling and recovery rules are natural elaborations of the Shiffrin (1970) rules.

D. SHORT-TERM STORE AND LONG-TERM STORE

The description of structure and retrieval given above is to a large degree independent of the theoretical assumption that memory is a two-phase system. Nevertheless, we find it useful for many reasons to place our retrieval system within a memory theory organized around a short-term store (STS) and long-term store (LTS). Atkinson and Shiffrin (1968) provide a prototype of such a system, but Shiffrin (1975) gives a more contemporary treatment. We shall review this system very briefly.

STS is postulated to be a temporarily activated subset of the information (and structure) in LTS, the permanent storage system. Sensory information presented to the system is analyzed automatically in a series of stages along many parallel paths. This analysis results in activation of information in LTS, and activation is equivalent to entry in STS. Alternatively, information is activated from LTS and placed in STS on the basis of internally generated probe cues, as described in earlier sections. Inevitably, both types of LTS activation occur together, so that sensory input (bottom-up processing) and information previously retrieved from LTS and presently still in STS (top-down processing) will jointly act to determine subsequent activation. In general, the activated information decays (becomes inactive) very rapidly, though small amounts of information may remain active in the absence of new input, or may be maintained in an active state for a long time through control operations like coding and rehearsal.

A major role of STS is its use as a working space for control processes of all sorts, including plans, coding, rehearsal, decisions, and so forth. The most important characteristic of STS is its limited capacity (see Shiffrin, 1976). There are limitations upon the rate of retrieval and examination of the contents of STS, upon the duration of residence in STS, upon the amount of information active in STS, upon the ability to focus and divide attention, and upon the rate of encoding of new information, among others.

These STS limitations affect retrieval in a number of ways. The impermanence and capacity limitations of STS limit the amount of information that may be sampled from the search set and maintained in an active mode. The limited rate of examination leads to sequential examination of one image at a time. The limitations on total STS load put bounds on the number of cues that may be used simultaneously (or on the sum of the weights).

E. LONG-TERM STORAGE AND LEARNING

Learning in our system consists of the formation of new associations, relations, and structures, in LTS, between elements and images already present in LTS, but concurrently active in STS. Although contiguity in STS may produce some storage in a passive mode, the strongest associations are formed when the subject gives attention to the material, and applies control processes like rehearsal, coding, evaluation, and relating the new material to already stored structures.

In most learning situations, storage will result from a combination of LTS retrieval and STS control operations. Indeed, since the coding of sensory input is also a form of LTS retrieval, all storage can be conceived as retrieval of a variety of LTS structures followed by the formation of new associative relationships between the retrieved structures.

The idea that the information simultaneously active in STS tends to be stored together is an extremely important aspect of the theory. In particular, it explains the prominence of temporal-contextual features (i.e., episodic memory—see Tulving, 1972) in memory images. Such temporal-contextual features include ''incidental'' information from the sensory environment and the subject's long-term store that happens to be present in STS at the time of a storage event. They might include the location, the temperature, the time of day, recent events, the subject's physical state, feelings, emotions, and recent thoughts. Each and every storage event will contain such temporal-contextual information to some degree, and this temporal context plays a prominent role in our retrieval theory and explanations of forgetting. In all retrieval situations, context will play a role as one of the probe cues, either by intent or accident. Presumably, the subject can, through attention, vary the weight assigned to this context cue, but such information will always be present in STS and will always play at least a small role as a retrieval cue. Whenever possible, of course, a knowledgeable subject will try to reinstate in STS as far as possible the contextual cues that had been present at the time that the to-be-recalled image had been stored.

F. LONG-TERM FORGETTING AS RETRIEVAL FAILURE

There are two basic reasons why an image may be retrieved better at time A than at time B. First, the cues utilized at time A may be more strongly associated to the image than those used at time B. Second, the strength or number of *other* images associated to the cues (even if the cues are the same) may be greater at time B than at time A. Everything else being equal, an increase of cue to image strength will increase both sampling and recovery probabilities [see Eqs. (1) and (2)]. On the other hand, for fixed cue to image strength, an increase in the strengths of cue to *other* images will reduce the sampling probabilities (though probably leaving recovery unaffected).

The increase in the strengths of association of cues to other images tends to be an inevitable consequence of new learning. This new learning will not necessarily lead to forgetting, however. The new information might be organized together or integrated with the old image so strongly that the retrieval of either set of information will then lead at once to retrieval of the other set. This integration could be conceptualized either as resulting in a single, new larger image, or as resulting in two tightly associated images. In the latter case, retrieval of one of the images could result in that image being used as a cue, and thereby eliciting the other image. This possibility is an example of a general principle: forgetting due to new learning occurs when the *same* cue is utilized in an attempt to locate one image among an increasing number of other images. On the other hand, the cues may be changed during the search so that each cue is related to a subset of the increasing number of images; in this event forgetting may be ameliorated or even reversed.

The decrease in the strengths of association of cues to image can be the result of several factors, chief of which is the change of context over time (see Estes, 1955; Bower, 1972). The context at the time of storage makes the best retrieval cue, but at the time of test, the context cue used may consist largely of the context information at the time of test, which will usually differ from the storage context by a greater amount as time between storage and test increases. Similar considerations apply to noncontextual cues, the general rule being that sampling and recovery will be worse as the retrieval cues chosen are less effectively associated to the desired image.

II. A Model for Free and Cued Recall

We develop the theory initially for the paradigm of free verbal recall. A list of N "unrelated" words is presented, one at a time. The presentation is sometimes followed by an arithmetic task to clear STS and restrict retrieval to LTS. The test involves recall of as many list words as possible, in any order. Usually, enough recall time is provided that the subject decides to cease retrieval, in the belief that memory is exhausted, before the recall period ends.

A. STORAGE ASSUMPTIONS

On the basis of coding and rehearsal operations that operate during list presentation, an LTS structure is generated and stored in LTS. The strength of associations of the cues at test to the LTS images is based on this structure.

Although many storage models are possible, we find it easiest to adopt the now traditional buffer rehearsal process of Atkinson and Shiffrin (1968). The buffer size is r. New words enter the buffer until it is full; then each new word replaces a randomly chosen word already in the buffer. The retrieval structure contains associative strengths between a general context cue and the images on the list, and between word cues and those images. It is assumed that these associative strengths grow linearly as a function of the total time that a word or a pair of words is rehearsed in the buffer. If we let t_i and t_{ij} be the times spent in the buffer respectively by I_i, and by I_i and I_j together then we assume: $S_{\mathrm{T}}(C, I_i) = at_i$; $S_{\mathrm{T}}(I_i, I_j) = S_{\mathrm{T}}(I_j, I_i) = bt_{ij}$, $t_{ij} \neq 0$; $S_{\mathrm{T}}(I_i, I_i) = ct_i$. Finally, even if two words on the list are not rehearsed together, they share context and are therefore assumed to have a nonnegligible residual retrieval strength, d: $S_{\mathrm{T}}(I_i, I_j) = S_{\mathrm{T}}(I_j, I_i) = d$, $t_{ij} = 0$. Thus the four parameters, a, b, c, and d, along with the buffer size r, completely determine the test matrix *at the start of retrieval.*

One additional storage process needs to be discussed. We assume that additional storage may take place during the course of retrieval itself. During retrieval, it will sometimes happen that a word is sampled, recovered, and recalled when a particular combination of cues is utilized as a probe set. Whenever this happens, but only when the recalled word is actually output, we assume that the strengths of the cues to the sampled image are incremented (i.e., increased). In addition, we assume that the self-association strength of the sampled image is also incremented in each such case. Thus we assume: $S'_{\mathrm{T}}(C, I_i) = S_{\mathrm{T}}(C, I_i) + e$; $S'_{\mathrm{T}}(I_i, I_j) = S'_{\mathrm{T}}(I_j, I_i) = S_{\mathrm{T}}(I_i, I_j) + f$; $S'_{\mathrm{T}}(I_i, I_i) = S_{\mathrm{T}}(I_i, I_i) + g$, where the primes indicate the strengths after incrementing, and e, f, and g are the parameters giving the amount of the increment in each case. As we shall see in the retrieval model, context will always be a cue, so the context strength and the self-association strength will always be incremented after a recall. The word-word strength will be incremented only in those cases in which a word was one of the cues.

B. RETRIEVAL ASSUMPTIONS

The heart of the retrieval model is Eqs. (1) and (2), giving sampling and recovery probabilities. In the present applications the weights W_i are set equal to 1.0.

At test time, any words still remaining in STS are output. Then retrieval from LTS begins. Figure 4 gives a flowchart corresponding to the first, main, phase of the model (which was written as a computer simulation).

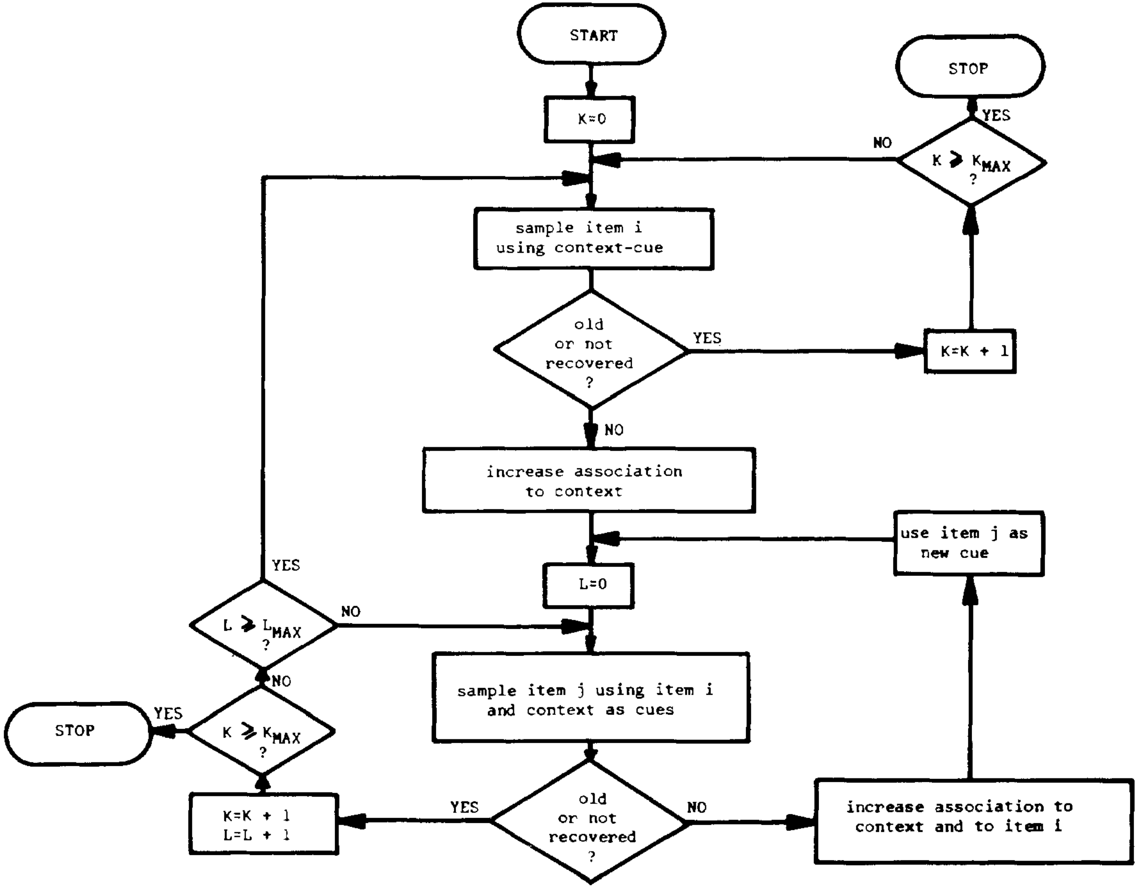


Fig. 4. A flowchart for phase one of the retrieval process in the computer simulation of SAM developed for free recall.

Search goes on until a criterion of K_{MAX} total failures is reached. A failure is every retrieval attempt that does not lead to recall of a new word. Thus, at the start of the recall from LTS the total failure counter K is set equal to 0. Next, the subject uses the general context cue, representing the context during study of that particular list, to sample from the images that are associated to that cue [Eq. (1) is used]. Suppose that image I_i has indeed been sampled; this is a failure if the image is ''old'' or if the image is new and cannot be recovered. An image is considered ''old'' if it has already been recalled, or if each of the present cues has previously been used to sample this image (unsuccessfully). This may be justified by the assumption that a given, fixed set of probe cues will always lead to recovery (i.e., activation) of the same set of features from a given sampled image. If the image is not ''old,'' the recovery rule of Eq. (2) is applied. If recovery fails, then a failure is counted, and context sampling is tried again. If recovery succeeds, then the strength $S_T(C, I_i)$ is incremented by e, and the self-association $S_T(I_i, I_i)$ is incremented by g.

After a successful recovery, the recalled word, I_i, is used as a retrieval cue along with context for the next sample [Eq. (1) is used]. Suppose image I_j is sampled. Then, as before, this is a failure if I_j is ''old'' or is new and cannot be recovered. (Note that in this case, the image may have been sampled unsuccessfully before, but the image will be considered ''new'' as long as the retrieval route $I_i - I_j$ is ''new.'') If I_j is new, then the recovery probability is determined by Eq. (2). If recovery fails, then a failure is counted, and L and K are increased by 1. If L_{MAX} and K_{MAX} are not reached, then the same cue combination is used again. If L_{MAX} is reached, then only context is used in the next cue set.

If recovery succeeds, then all relevant strengths are incremented: $S_T(C, I_j)$ is incremented by e, $S_T(I_j, I_j)$ is incremented by g, and $S_T(I_i, I_j)$ is incremented by f [as is $S_T(I_j, I_i)$ since we assume bidirectionality].

If I_j has been recovered, then this word is used as a cue, along with context, in the next cue set. This entire process continues until K_{MAX} total failures are reached.

In summary, extensive use is made of interitem associative routes: whenever a new word is recalled it is used as a cue either until L_{MAX} failures accumulate or until a new word is recalled, in which case the new word is used as a cue. Of course, it could be argued that all interitem routes have not been fully explored, since a switch to a new word cue may occur before search with the previous word cue has been exhausted.

For this reason a final ''rechecking'' process is incorporated in the model after the K_{MAX} criterion has been reached. Every word that has been recalled (presumably they are written down and hence available) is used as a cue, along with context. L_{MAX} samples are made with each such cue combination. Any new words recalled during this period are also "rechecked." With this rechecking process added, it may be argued that the subject feels all retrieval routes have been tried and exhausted.

It is not too difficult to see that this retrieval model is at least potentially able to explain a variety of data in free recall, especially if one recognizes that the model combines features of two powerful models, namely, the model of Shiffrin (1970) and the FRAN model of Anderson (1972).

C. PARAMETERS

The model described above is ready to be applied to the data from free recall studies. The parameters are a (context to image strength), b (image to image strength), c (image to self strength), d (residual strength), e (context to image increment), f (image to image increment), g (self increment), K_{MAX} (total failure stopping criterion), L_{MAX} (stopping criterion for a word cue), and r (buffer size).

At first glance, 10 parameters seems quite a high number, even though we shall fit a great deal of data from a variety of paradigms. For example, Shiffrin (1970) fit a great deal of free recall data with just three parameters. This objection is ameliorated by the following factors. We can show that most of the present parameters, and their precise values, are *not* essential for the fit of the model to most of the data. The parameters are listed above for generality, even though some are never varied and others are equated before fits to the data are begun. Some of the parameters are given nonzero values and included in the fit merely to demonstrate that the presence of the processes they represent will not harm the ability of the model to predict the data. In fact, we have set many of these parameters to zero, and no harm to the model's predictions results. However, each of these parameters represents processes that we feel are needed on logical grounds, or needed to deal with data from at least one of the studies to be discussed in this paper. The roles played by the various parameters have been extensively explored by simulation means, as have certain process assumptions, and the results of these explorations will be summarized briefly or reported in detail in the remainder of the article.

III. Applications of the Theory

A. FREE RECALL: SERIAL POSITION, LIST LENGTH, PRESENTATION TIME

Primacy and recency effects are predicted by our model as a consequence of the buffer assumption. These effects are therefore easy to predict but they are not very informative concerning the LTS retrieval process. A more interesting result is that the model is able to describe the serial position curves for different list lengths and presentation times with the same set of parameter values. Figures 5, 6, and 7 show the serial position curves obtained by Murdock (1962) in a task including STS recall. In Murdock's experiment six groups of subjects each had a different combination of list length and presentation rate. The six conditions were 10–2, 20–1, 15–2, 30–1, 20–2, and 40–1, where the first number refers to the list length and second number indicates the number of seconds that an item was presented.

The parameters of our model were very roughly estimated by a Monte Carlo simulation technique from the data of conditions 10–2, 20–1, 20–2, and 40–1. These parameter estimates should not be regarded as optimal since only a limited search of the parameter space was feasible. Moreover, the parameter space is quite shallow, so that many other combinations of parameter values will give a fit about equally good.

Many of the parameters were set arbitrarily, rather than estimated. The buffer size, r, was set equal to 4, and K_{MAX} set equal to 30, on the basis of previous work (Shiffrin, 1970). L_{MAX} was set equal to 3 (a value that later simulations showed produced near maximum recall). The values of e, f, and g (incrementing) were set equal, d (the residual) was set to one fifth of b, and then a, b, and e were estimated. The values that gave a "best" fit, roughly, were $a = .055$; $b = .02$; $e = .6$. The resulting

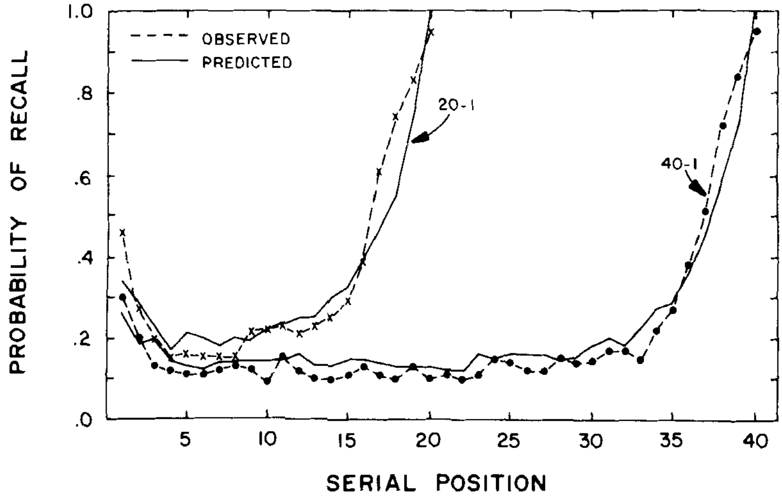


Fig. 5. Serial position results (Murdock, 1962) for 20- and 40-word lists at 1 sec/word. Predictions from the SAM model with $r=4$, $K_{\mathrm{MAX}}=30$, $a=c=.055$, $b=.02$, $d=.004$, $e=f=g=.6$. These parameters are also used to derive the predictions in Figs. 6 and 7.

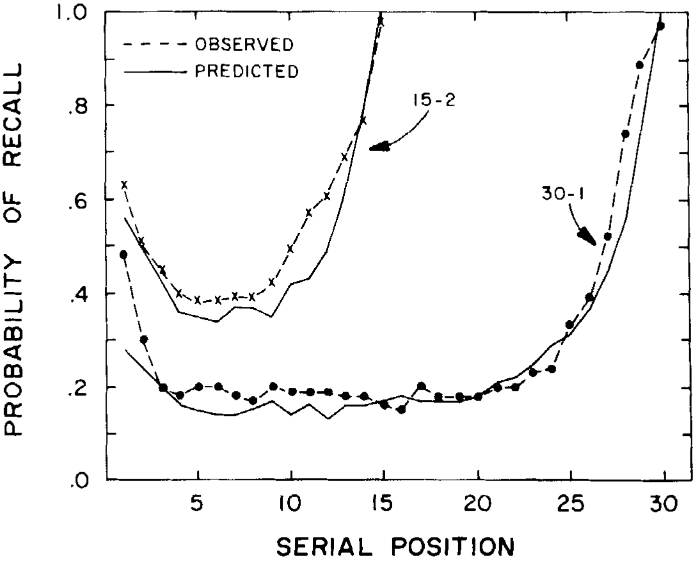


Fig. 6. Serial position results and SAM's predictions for a 30-word list at 1 sec/word and a 15-word list at 2 sec/word. (Data from Murdock, 1962.)

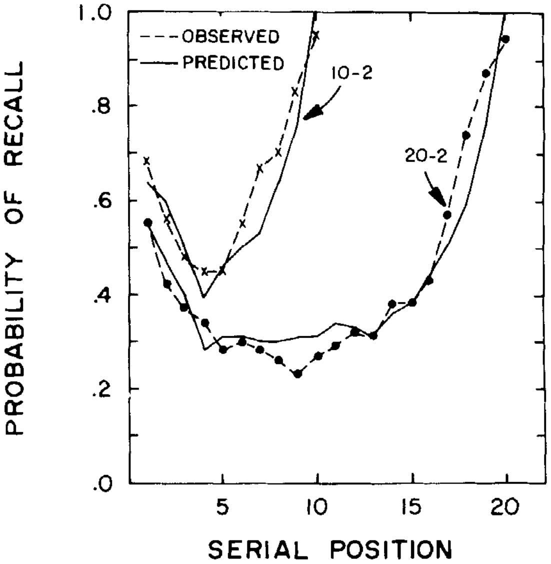


Fig. 7. Serial position results and SAM's predictions for 10- and 20-word lists at 2 sec/word. (Data from Murdock, 1962.)

predictions are shown in Figs. 5, 6, and 7. Clearly, the fit is quite adequate.

Note that both the predictions and the data show a list-length effect: the probability of recall is a decreasing function of list length. This list-length effect is predicted by the model because the search termination criterion is exceeded sooner for the longer lists, relative to the list length, i.e., *relatively* fewer samples are made from a longer list than from a shorter list. The probability of sampling an item is therefore lower for an item from a longer list. This effect is predicted by the model even when the criterion is set very high ($K_{\mathrm{MAX}} = 100$) or when a stop-rule of K_{MAX} *consecutive* failures is used. Thus, this prediction is a consequence of the basic structure of the model: a sampling-with-replacement retrieval process coupled with a fixed termination criterion (i.e., the criterion does not vary with list length).

We should note that many of the process assumptions and parameter values are not essential for predicting the Murdock data. If rechecking is eliminated, a *very* slight adjustment in the *a, b,* and *e* values will produce an equivalent fit. If the residual association, *d,* is removed (set to 0), an equivalent fit is obtainable by changing *a* to .065 and *b* to .015. If the stopping rule is changed to K_{MAX} *consecutive* failures, an equivalent fit is obtained without changing any parameters, but letting $K_{\mathrm{MAX}} = 15$ (including the three failures in the last search with a word cue).

Although the fit to Murdock's data is quite good, list length and presentation time per item were not varied over a very wide range. Roberts (1972) reported the results of a large, well controlled study, where four list lengths (10, 20, 30, or 40 items) and five presentation rates (.5, 1, 2, 4, or 8 sec per item) were varied in a factorial design. His results are shown in Figs. 8 and 9 in the top panels, where we have averaged the data for the auditory and the visual presentation modes. Note that these results include recall from short-term store since no interpolated task was given. These results show that the mean number of words recalled is not a linear function of the total presentation time (Murdock, 1960) but a negatively accelerated function as found by Waugh (1967). They also clearly show that the total-time hypothesis (Murdock, 1960) is incorrect: equal total presentation times do not yield equal levels of recall.

Figures 8 and 9 also show the predictions derived from the present model. These predictions include of course the recall from the STS-buffer. Parameters were estimated as in the case of Murdock's data: *a, b,* and *e* were estimated. The best fitting values were $a = .10$, $b = .10$, $e = .70$. The quality of the fit to the data is seen most easily in Fig. 10, which compares probabilities of predicted and observed recall for each of the 20 points in Fig. 8 (or 9). Obviously the fit of the model is quite

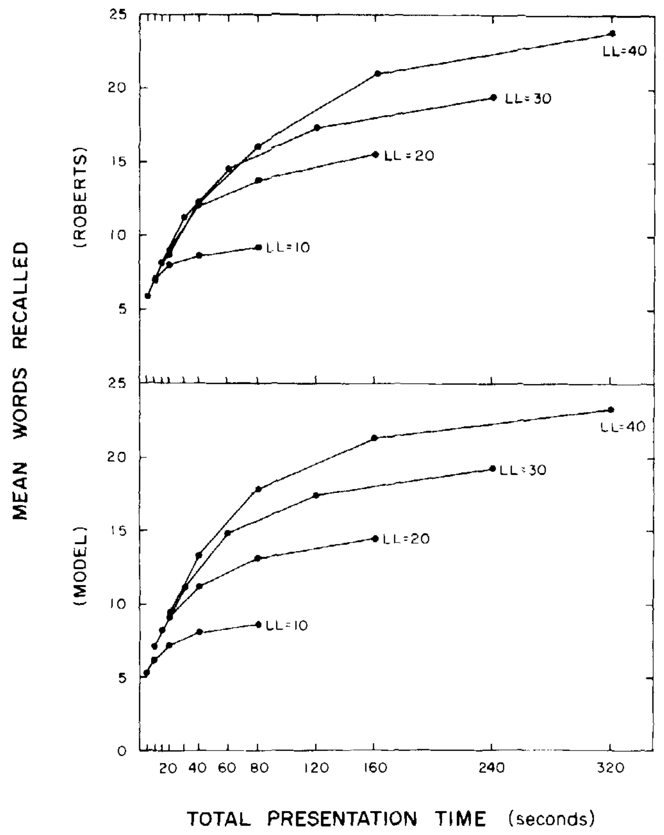


Fig. 8. Observed data (Roberts, 1972—top panel) and predictions of SAM (lower panel), for mean words free-recalled as a function of presentation time and list length (LL). The parameters of the model (see text): $r=4$; $K_{\mathrm{MAX}}=30$; $L_{\mathrm{MAX}}=3$; $a=c=.10$; $b=.10$; $d=.02$; $e=f=g=.70$.

satisfactory. Note that in our model presentation time per item has its effect mostly on the probability of recovery, not on the probability of sampling [see Eqs. (1) and (2)]. There is only a small effect on the probability of sampling due to the fact that the increment upon successful recovery is the same constant in all conditions and thus relatively higher in the case of a lower presentation time per item. As with Murdock's data, a list-length effect is evident in Fig. 8. Our model predicts such effects because relatively fewer samples are made from a longer list. Thus, the list-length effect is predicted to be a retrieval effect, not a storage effect. Of course, the subjects do not know how long the list is going to be.

As with Murdock's fit, certain processing assumptions and parameter values are not essential in this case. Rechecking makes little difference. Setting $d = 0$ can be compensated for by raising a to .12, with equally good results. Changing to a K_{MAX} consecutive failure rule, with $K_{\mathrm{MAX}} = 11$, and the other parameters unchanged, gives a fairly good fit, but with the predicted points in Fig. 10 lying along a line with slightly greater slope than the observed points. This is easily fixed, however, by slight changes in the values of the other parameters. All in all, it seems clear that the predictions of the list-length effects and the effects of presentation time are the result of the basic structure of the model and not of the particular parameter values or assumptions used.

In the applications both to Murdock's data and Robert's data, no mention has been made of the effects of the interword-association retrieval routes. The reason is simple. The word cue searches are not needed to

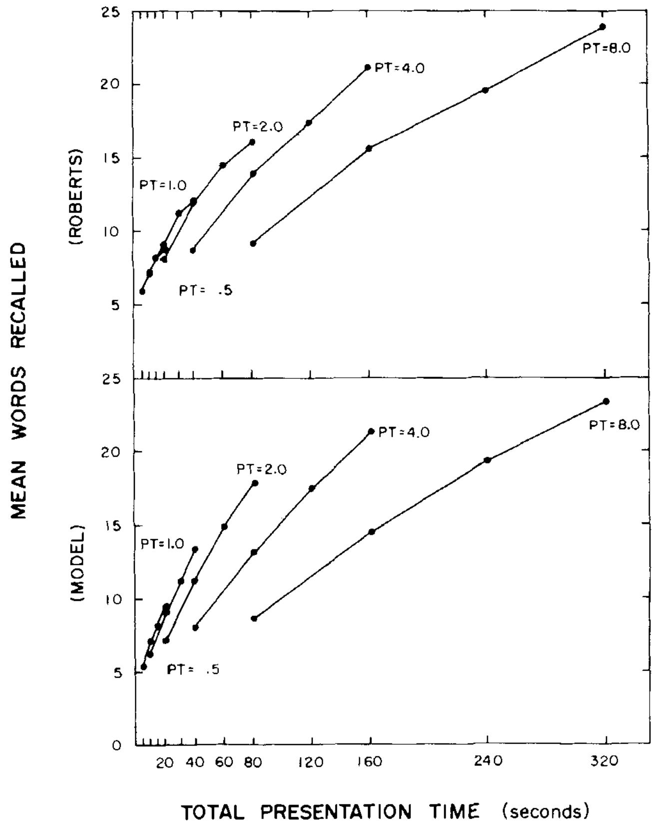


Fig. 9. Same as Fig. 8, except parameter of each curve is presentation time per item.

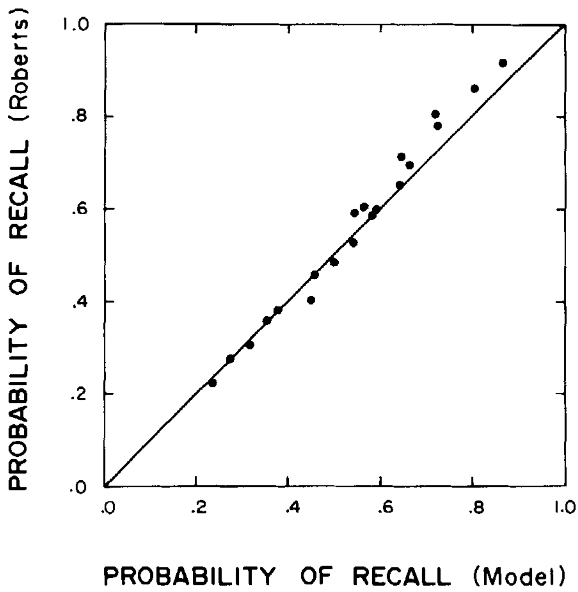


Fig. 10. Comparison of the predicted and observed probabilities of recall, corresponding to each of the 20 points in Fig. 8 (or 9).

predict these data (see Shiffrin, 1970). However, the word-cue searches are crucial for many of the applications to be covered later, and it is noteworthy that the inclusion of such a process does not affect the predictions. Interestingly enough, we have even found that virtually identical predictions, for the same parameter values, can be obtained if it is assumed that the *two* most recently recalled words are used as cues along with context (until L_{MAX} failures accumulate). Such results suggest that a subject may not have a uniformly optimal strategy of cue selection. They further suggest that different subjects, or the same subject at different times, may use different numbers of item cues without much affecting total recall.

Finally, consider the strictness of the search stopping criterion (the value of K_{MAX}). The fact that K_{MAX} was not estimated and yet a good fit was obtained suggests that the criterion may be shifted, one or two other parameter values changed slightly, and equally good predictions obtained. This is, in fact, the case. More important, it is one of the great successes of the model that empirical manipulations designed to change the stopping rule produce results that are extremely well predicted by the model, with the only alterations occurring in the value of K_{MAX}. We turn next to such studies and predictions.

B. EXTENDED RECALL, REPEATED RECALL, HYPERMNESIA, AND INTERRESPONSE TIMES

A model for free recall should exhibit several closely related properties: (1) the criterion for cessation of search should be such that a reasonable subject could be expected to "give up" at that point; (2) the temporal point at which time-unlimited search stops should be a point at which few new items are being recalled (for all conditions); (3) predicted cumulative recall functions, at times before subjects cease searching, should grow at a rate similar to that seen in the data; (4) if subjects are induced to extend their search beyond the point of normal cessation, the number of additional words recalled should be predictable by shifting the stopping criterion, or by otherwise altering the model's stopping strategy in a manner consistent with the instructions and task demands.

Consider first the stopping rule. When a *total failure* rule is used, even with a high value of K_{MAX}, it is not obvious that new recalls will be occurring at a slow rate just before search ceases. In fact, however, even with $K_{\text{MAX}} = 30$, the output rate is quite low most of the time when search stops. This is supported by the observation that in the models for the Murdock and Roberts data, *consecutive failure* rules of 15 and 11, respectively, give predictions virtually equivalent to those for the total failure rule. Certainly it seems reasonable that a subject should cease recall after such a long string of failures (ignore for the moment the fact that we assume rechecking to take place *after* this criterion is reached).

In order to gain a clearer picture of the output rates, we give in Fig. 11

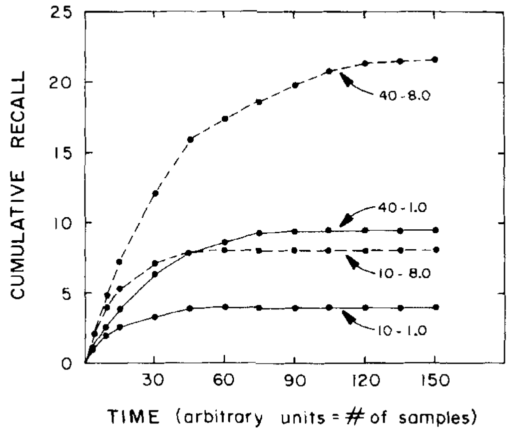


Fig. 11. Predicted cumulative output functions for list lengths of 10 and 40 and presentation times of 1.0 and 8.0 sec per item. SAM's parameter values are the same as for Fig. 8, but note that STS retrieval has been deleted.

predicted cumulative output functions for four of Robert's conditions, with Robert's parameters. Note that a stop rule followed by rechecking is assumed to apply to each "subject"; the cumulative functions simply give total cumulated recall over subjects divided by the number of subjects, so that after a long "time" most of the "subjects" have ceased trying to recall, and only a few of the subjects are contributing new recalls to the cumulative functions. This explains why the lower functions in the figure reach a nongrowing asymptote—all subjects have stopped retrieving. These predicted functions show a very important property; the rate of approach to a higher final asymptote is slower. In fact, a considerable literature attests to just this fact (see Johnson, Johnson & Mark, 1951; Bousfield, Sedgewick, & Cohen, 1954; Indow & Togano, 1970). Note that predicted recall grows at a reasonable rate for quite a long time when the list length and presentation time per item are large. What should be the most reasonable stopping rule in such a case is difficult to judge.

The predictions in Fig. 11 show what happens when subjects are assumed to use a normal stopping rule, so that search ceases relatively quickly. It might be asked, what are the predicted cumulative output functions if subjects are induced to search for very lengthy periods without stopping. Typical predictions are shown in Fig. 12. The Robert's parameters are used, except there is no stopping rule at all. List length is set to 40, presentation time to 4 sec/word. The dashed curve is an extreme case in which *no* rechecking is assumed, so that almost all samples late in retrieval use context only as a cue (only after a new recall occurs is there a brief period, with criterion = $L_{\text{MAX}} = 3$, of cuing with word + context). The solid curve gives an extreme case in which rechecking occurs whenever a multiple of 50 samples occurs (unless rechecking is still underway at that point). Early in search, for this list length and rate, rechecking actually harms recall (see the portion of the curves between 50 and 100). This occurs because rechecking gives rise to new words very slowly. On the other hand, rechecking ensures that new retrieval routes become available. That is, an image may have been sampled but not recovered with context and words $i_1, i_2, \ldots, i_M$ as cues; this image can still be recovered if sampled with a new cue, word i_{M+1}. Thus the rechecking curve continues to rise, albeit slowly, until it surpasses the other curve. This is a general property of the rechecking assumption: it causes cumulative output functions to continue to grow for longer periods of time. Finally, it may be interesting to compare the levels of recall after 500 samples under either assumption, about 29 words, with the predicted level if a total failure rule of 30 is used (see Fig. 8), 21.4 words. For the rechecking curve, such a level corresponds to about 125 samples. Clearly, more words are predicted to be available in memory when search

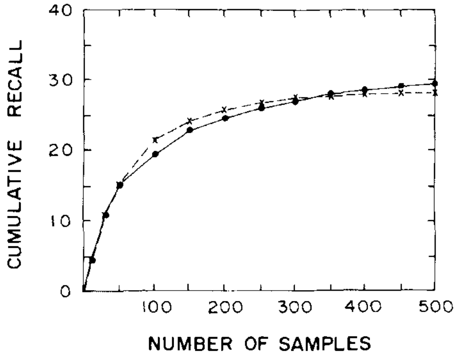


Fig. 12. Cumulative output predictions for SAM without a stopping rule. Dashed curve: no rechecking assumed. Solid curve: rechecking assumed every 50 samples. Parameter values as in Fig. 8; list length = 40, presentation time = 4 sec/word.

stops with $K_{MAX} = 30$, but an enormous effort may be required to retrieve them.

We might now ask whether cumulative *data* functions show any of these properties. Figure 13 shows cumulative functions from Roediger and Thorpe (1978) who induced subjects to continue to try to recall for 21 min, corresponding to the assumptions used in Fig. 12. Without attempting to estimate parameters, is seems clear that real subjects show increases in recall over quite long periods of time, and that the growth functions are quite similar in form to SAM's predictions.

Figure 13 also shows what happens when the subject is given three consecutive recall periods of 7 min each, the subject beginning over in each new recall period. The cumulative curves shown for this case ignore any multiply recalled words and simply count new words recalled. On the other hand, if one counts total words recalled during each 7 min period, then this total increases in each period, especially for pictures, as shown in the top left panel of Fig. 14. This phenomenon has been called "hypermnesia" by Erdelyi and his colleagues (see Erdelyi & Kleinbard, 1978) and interpreted as some sort of "negative forgetting." Roediger and Thorpe's data shown in Figs. 13 and 14 seem to make it clear that the effect is merely a consequence of more total time available for recall, along with a result that fewer previously recalled items are forgotten than new words are recalled.

It is not immediately obvious that SAM should predict this "hypermnesia" result, despite the predicted growth in cumulative output, since it is difficult to judge intuitively how many previously recalled words will be predicted to be forgotten in a following recall period. It is easy to apply SAM to this task, however. Since Roediger and Thorpe

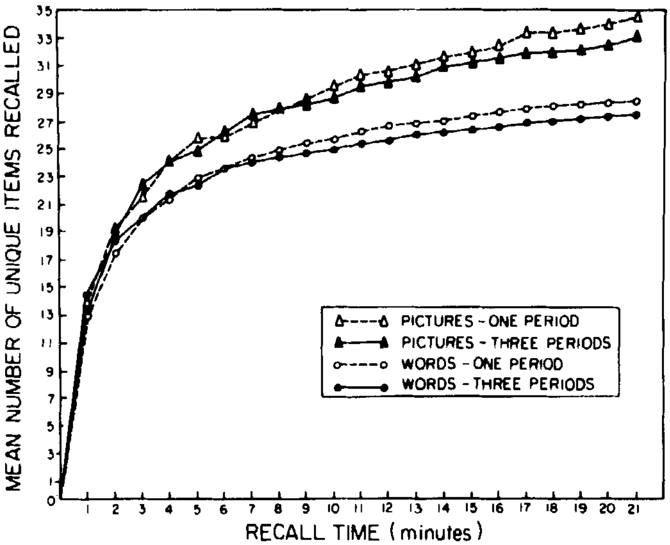


Fig. 13. Mean cumulative recall of unique items for subjects presented pictures or words and given either three 7-min tests or one 21-min test (taken from Roediger & Thorpe, 1978).

(1978) use long recall periods (7 min) and report that the hypermnesia effect is obtained only in this case, we deleted the stop rule from SAM and simply assumed that 150 samples were made in each recall period. Rechecking of all previously recalled items within the current recall period was assumed after each 50 samples. List length was set equal to 40, presentation time per item to 1 sec, and parameters chosen to represent high strengths ($a = .2$, $b = .1$, $e = .7$) or low strengths ($a = .1$, $b = .05$, $e = .7$). The other parameters were those for Roberts data ($r = 4$, $L_{MAX} = 3$). The predictions are shown in the middle upper panel of Fig. 14. Under either high or low strength assumptions, the increase across successive recall periods is observed. The upper right-hand panel shows similar predictions when the incrementing parameter is greatly increased, to 3.0.

In order to determine which features of the model are responsible for the "hypermnesia" prediction, several alternate assumptions were used. The "normal" version already discussed assumes that an image that has previously been sampled but not recovered may still be recovered later if the cue-set contains at least one cue that is new for that image. This assumption is denoted "alternate retrieval routes." The lower left-hand panel shows that the hypermnesia prediction is reduced but not eliminated if the "alternate routes" possibility is eliminated. In this case, only one recovery chance is possible for a given image, but incrementing remains.

The center lower panel shows that hypermnesia is also predicted if incrementing is set to .0, but alternate routes remain. However, if *both* "alternate retrieval routes" and "incrementing" are removed from SAM, then it may be shown that no change in recall is predicted for successive recall periods. Finally, the right-hand lower panel shows that the amount of increase is lowered but not eliminated if increments are allowed to take place anew in each successive recall period (thus an item already incremented to a cue, can receive another increment to that cue in a later recall period).

One or two final points should be mentioned about the "hypermnesia" prediction. First, if the number of samples per recall period is reduced, the predicted increase in recall lessens considerably. Such a prediction accords with data reported by Tulving (1967) and Donaldson (1971) (although these results are difficult to interpret because the first recall

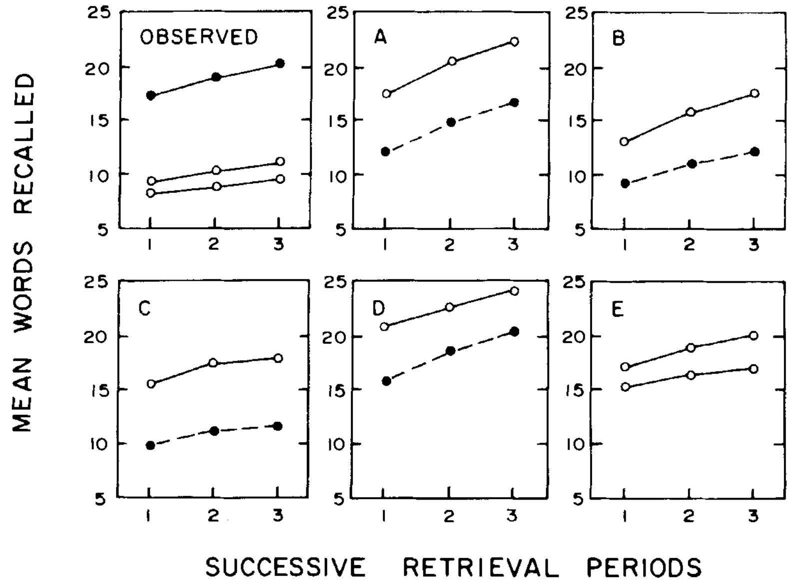


Fig. 14. Mean number of words recalled in each of three successive retrieval periods for the same list. Top left panel: observed data from Roediger and Thorpe (1978). Panels A to E give predictions. (A) Alternate retrieval routes assumed: solid curve: $a=c=.2$, $b=.1$, $d=.02$, $e=f=g=.7$; dashed curve: $a=c=.1$, $b=.05$, $d=.01$, $e=f=g=.7$. (B) Alternate retrieval routes assumed: solid curve: $a=c=.2$, $b=.1$, $d=.02$, $e=f=g=3.0$; dashed curve: $a=c=.1$, $b=.05$, $d=.01$, $e=f=g=3.0$. (C) No alternate routes: solid curve: $a=c=.2$, $b=.1$, $d=.02$, $e=f=g=.7$; dashed curve: $a=c=.1$, $b=.05$, $d=.01$, $e=f=g=.7$. (D) Alternate routes assumed: solid curve: $a=c=.2$, $b=.1$, $d=.02$, $e=f=g=0$; dashed curve: $a=c=.1$, $b=.05$, $d=.01$, $e=f=g=0$. (E) Increments apply again each new retrieval period; top curve: alternate routes assumed: $a=c=.2$, $b=.1$, $d=.02$, $e=f=g=.7$; bottom curve: no alternate routes: $a=c=.2$, $b=.1$, $d=.02$, $e=f=g=.7$.

includes an STS component). On the other hand, if the number of samples remains high, but rechecking is eliminated, a large increase across recall periods is still predicted. Finally, if the assumptions are changed so that each recall period ends when a stopping criterion is reached, a fairly large "hypermnesia" effect is still predicted (for either type of stopping rule).

The various effects and predictions of this section have all been concerned with cumulative recall over time. We now wish to look at a finer-grained temporal aspect of recall, interresponse times (IRTs). The basic data of interest were collected by Murdock and Okada (1970). Each of 72 subjects was given 20 lists of 20 words each to free recall. Words were presented visually, and the free recall was spoken and tape recorded. Two presentation rates were used, 1 word/sec or 2 words/sec, but these were collapsed together in the reported analyses. Figure 15 shows the mean interresponse time between each of the consecutive ordinal output positions, partitioned separately for each different number of total words output. Because there were insufficient data when fewer than four or greater than nine words were recalled, these curves are not shown.

We did not attempt to fit the exact data of Murdock and Okada (1970), since their data include a STS-component. Simulation of their experiment would therefore necessitate the prediction of retrieval rates from STS. Because our main interest is on retrieval from LTS, this did not seem to

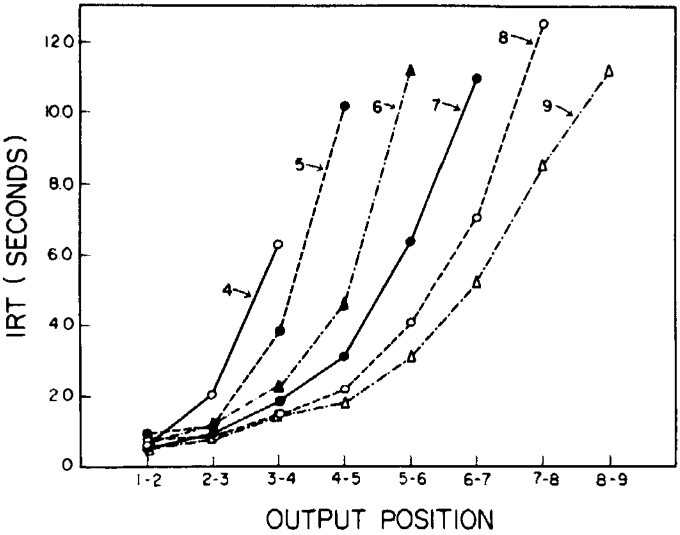


Fig. 15. Mean interresponse time (IRT) as a function of ordinal position in output. Data partitioned according to total number of words recalled (4–9). (From Murdock & Okada, 1970.) Data include STS retrieval.

be desirable. We set list length equal to 15 and presentation rate equal to 2 sec per item. The same model was used as for Roberts' data, except that the STS-buffer was cleared before recall began, and the rechecking assumptions are slightly altered. Since the subjects in this study used spoken, not written, recall, they would not have all previously recalled items available. It was therefore assumed that after the K_{MAX} criterion had been reached, the "subject" continues the search via a rechecking process which uses as cues items recovered subsequent to the K_{MAX} point, even if these items had already been recalled. It was assumed that the number of failures in this second phase of search, K_{MAX2}, would be set equal to the total recall in the first phase multiplied by L_{MAX}. This assumption makes the rechecking effort similar to that assumed for earlier versions of the model. In phase two, search begins with the context cue, and as soon as *any* recoverable item is sampled, switches to it as a cue along with context. Then this cue set is used until a new recall occurs, in which case a switch is made to the new item as a cue, or until L_{MAX} failures accumulate. In this case, the context cue alone is used in the successive samples, until a recoverable item is reached. This process continues until K_{MAX2} failures are reached.

Figure 16 shows the predictions, based on 5000 simulation runs with Robert's parameters. The predicted curves are very similar to the empirical curves obtained by Murdock and Okada (1970). Several features that were noted by Murdock and Okada are also evident in the simulated data. First of all, the interresponse times increase in a positively accelerated fashion as recall proceeds. Second, for a fixed output position, the interresponse times were shorter the more words there were yet to recall. Finally, at any given output position the interresponse time is a good predictor of the number of words yet to recall. Of course these predictions are based on the particular rechecking assumptions that have been made, and a different set of assumptions would undoubtedly shift the predictions (for example, without *any* rechecking or a comparable process the predicted curves are almost linear, with decreasing slopes toward the right of the figure). The lesson from this simulation is simply that IRT results like those of Murdock and Okada are quite consistent with a SAM-like model.

C. CATEGORIZED FREE RECALL: CUING, OUTPUT INTERFERENCE, TEST ORDER

In categorized free recall the list of words that is presented to a subject is divided into a number of conceptual categories (e.g., four-footed animals, professions, tools, etc.). The words belonging to a particular category may be presented contiguously (blocked presentation) or in random

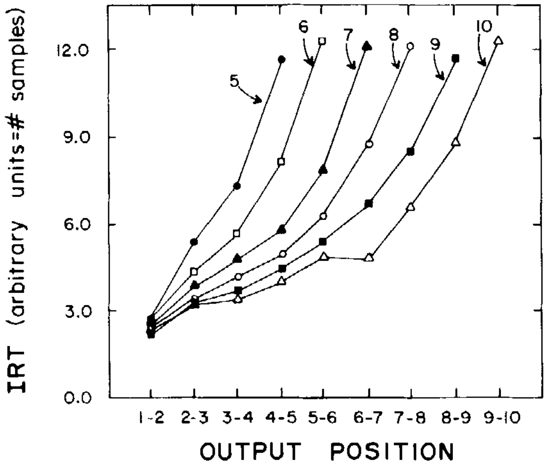


Fig. 16. Predicted mean interresponse times partitioned according to the total number of words recalled (5-10); STS retrieval deleted from predictions. Parameters as in Fig. 8; list length = 15; presentation time = 2 sec/word; special model assumptions given in text.

list positions (random presentation). The present theory will be applied to a blocked presentation procedure only. (In any event, the results mentioned by Cohen, 1966, suggest that if subjects are aware of the categorical nature of the list there are no qualitative differences in the gross results between blocked and random presentation.)

In principle, extending the SAM model to the categorized situation is quite simple. Just as is the case for context information, it is assumed that category information is stored as part of each image, and that category information may be used as a cue. Let us denote an item in a cued category by using a prefix c, and an item in a different category than that of the cue by using a prefix nc. Then the SAM model sets the strength of category cue to c-image to be a linear function of rehearsal time, and the strength of category cue to nc-image to be equal to a residual value. Furthermore, there are separate increments for the category to c-image strength, and the category to nc-image strength, when either of these is recovered and output in the presence of the category cue. Such a model has been fit quite successfully to a variety of data from categorized paradigms.

We have found, however, that not all this machinery is needed to capture the essence of the mechanisms of the categorized situation. A much less powerful model is quite capable of predicting almost all the effects. This simpler model ignores all interitem strengths and retrieval routes (i.e., items cues are not used). Instead, only context and category cues need be considered. It is this model that shall be presented and utilized in the following sections, since it illustrates the main points without confusing the issue.

Since we are not concerned with serial presentation position effects in the categorized task, it is also much simpler to replace the buffer storage system with a fixed strength assumption. In particular, the context to image strength is set equal to a times the presentation time for that item, and the category to c-image strength is set equal to B times the presentation time. The category to nc-image residual strength is set equal to D. The increment for context-to-image strength was set equal to e and the increment for category to c-image strength was also set equal to e.

An excellent demonstration of the power of even this simplified model to explain the results of categorized studies may be obtained by applying it to a well controlled study by Tulving and Pearlstone (1966). They varied three independent variables: (1) list length—12, 24, or 48 list items, (b) number of words or items per category—1, 2, or 4 items per category, and (3) type of recall test—either a cued or a noncued recall test. In noncued recall the subjects were given a standard free recall instruction, i.e., they were told to write down all the words they could remember as having been on the list. In cued recall the subjects were given a list of all the category names and then tried to recall as many words as possible. In this experiment the members of each category were presented in a blocked fashion, preceded by the category name. Subjects were instructed carefully that they were to remember only the category members, not the category names. The presentation time was 1 sec for each item and 3 sec for each category name. The amount of recall time given was proportional to the list length (1 min for every 12 items). Following the first recall test all subjects were given a second recall test. This second test was always a cued recall test. The results for the first test are given in Figs. 17 and 18 as the solid points.

The application of the model to this data is fairly straightforward. In the case of cued recall it is assumed that each category cue is used until a criterion of L_{MAX} total failures is reached (no rechecking). The items are sampled using both the context and the category cue. The probability of sampling is therefore proportional to the product of the item-to-context associative strength and the associative strength between that item and the category that is tested. The probability of recovery is given by the usual exponential transformation of the sum of these two associative strengths. Retrieval of an item outside the category being tested is assumed to be a failure, on the reasonable basis that a subject always recognized whether an item belongs to the category being tested. The contextual and category associative strengths are incremented upon successful recall of an item.

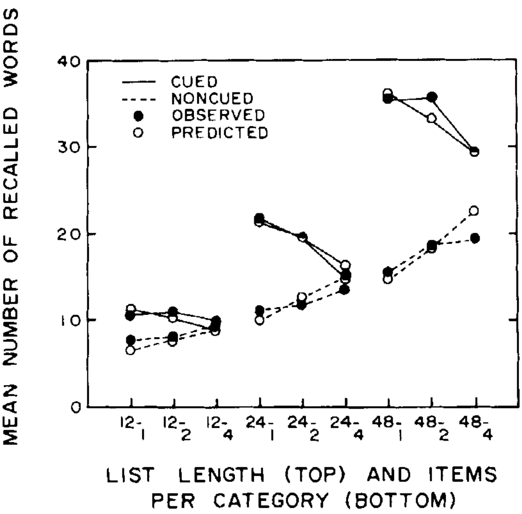


Fig. 17. Mean number of words recalled in the first recall test as a function of list length and words per category, for testing with or without category names provided. (Data from Tulving & Pearlstone, 1966.) Predictions from simplified SAM model, described in text.

Note that due to the assumption of a nonzero residual strength of the association between a category and an item belonging to a different category the model predicts an effect of the number of, and the contextual associative strength of, the items belonging to other categories.

In the case of noncued recall it is assumed that the subject first samples one of the items using only the context cue. Upon successful retrieval of an item the subject will generate (with probability 1.0) the category name of which that word is a member. Both the context and the category strengths are incremented. Next, the subject tries to recall items from within that category until he reaches a criterion of L_{MAX} failures. Contextual cues as well as the category cue are used in this restricted search. As before, retrieval of an item outside the category tested is counted as a failure. Thus, this category search is exactly the same as in the case of cued recall. When the criterion of L_{MAX} failures has been reached the subject discards the category cue and continues sampling using only the context cue. This goes on until a criterion of K_{MAX} total failures with the *context* cue has been reached. Note that failures that are made during category searches are not counted as part of these K_{MAX} failures. Thus, recall stops when the subject believes that he can find no more new categories.

The above model for categorized free recall is similar to the model proposed by Shiffrin (1970) and Rundus (1973). However, neither of these models allows for sampling items outside the cued category. Therefore, they have difficulty explaining total list-length effects upon within-category recall, and order effects of testing successive categories in cued recall (e.g., Smith, 1971; Roediger, 1973).

The model was fit to the data in Figs. 17 and 18 and a good fit was obtained for the following parameter values: $a = 1.2$; $B = 1.7$; $D = 0.2$; $e = 2.0$; $L_{\mathrm{MAX}} = 15$; $K_{\mathrm{MAX}} = 20$. The predictions are also given in Figs. 17 and 18.

Inspection of Figs. 17 and 18 reveals that the advantage of cuing with the category name decreases with increasing category size (for constant list length) and increases with increasing list length (for constant category size). The model predicts these effects because with increasing category size relatively more categories are accessed in noncued recall, thereby eliminating the advantage of the cued group. With increasing list length, however, relatively fewer categories are accessed in noncued recall, which increases the advantage of cuing.

These data were analyzed by Tulving and Pearlstone (1966) in terms of the two response measures used earlier by Cohen (1963); category recall (R_c), the number of categories of which at least one member was recalled, and words-within-category recall ($R_{w/c}$), the ratio of the total number of words recalled to the number of categories recalled. Thus, a

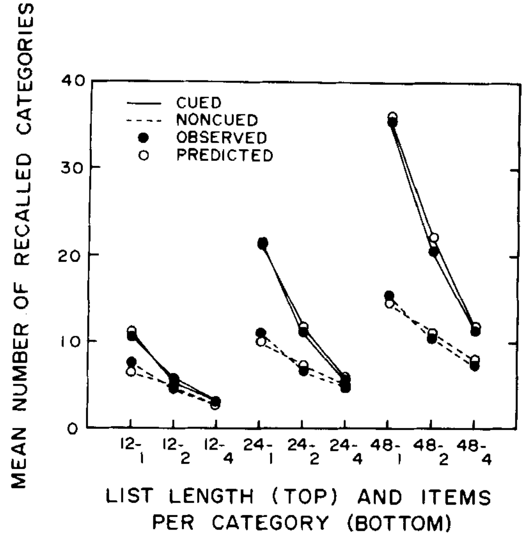


Fig. 18. Mean number of categories recalled in the first recall test as a function of list length and words per category, for testing with or without category names provided. (Data from Tulving & Pearlstone, 1966.) Predictions from simplied SAM model, described in text.

category is not considered to be "recalled" when no member of that category is recalled. Analyzed in this way the data show that the probability of recalling a category was higher for the cued group, that this probability decreased with increasing list length in both the cued recall and the noncued recall condition, and increased with increasing category size in both conditions. Naturally (judging by our fit), when the predicted data are analyzed in the same way, the same effects are obtained. The reasons why the model predicts these results are evident, if we keep in mind that sampling *within* a category is predicted to depend upon the number and strengths of items in *other* categories, due to the residual associations, D. Thus, for example, the probability of recalling an item from a cued category will go down if there are more items in other categories on the list.

In seeming conflict with this reasoning is the finding that words recalled per category having at least one recalled member did not vary much with total list length (in both the data and predictions). However, this seeming paradox disappears when the data are reanalyzed in a nonconditional fashion. The possibility of *no* recalled members from a cued category must be taken into account. Figure 19 gives the observed data and the predictions for items per cued category (with unconditional scoring). Clearly the list length effect is present in both data and predictions.

One other interesting finding observed by Tulving and Pearlstone (1966) for the first test is that the probability of recalling a member of a category, given that at least one member was recalled, was a decreasing

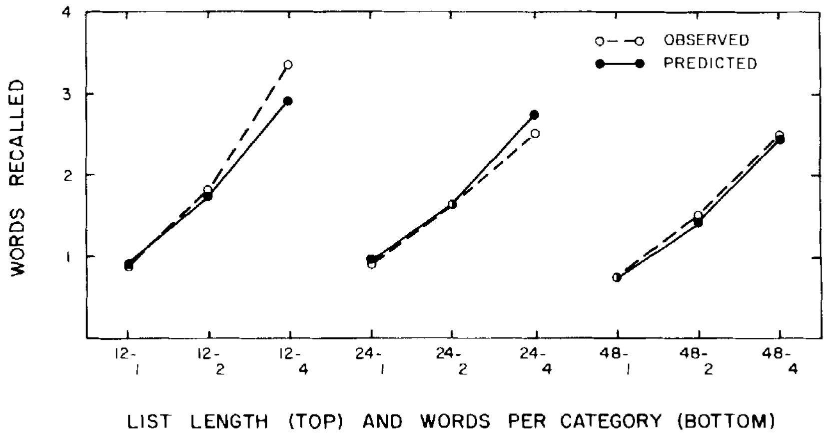


Fig. 19. Predicted and observed mean number of words recalled per cued category in test one as a function of list length and words per category. (Data from Tulving & Pearlstone, 1966.) Predictions from simplified SAM model, described in text.

function of category size. This is of course similar to the usual list-length effect in (uncategorized) free recall, and this result is predicted by our model for exactly the same reasons.

As was mentioned above, Tulving and Pearlstone (1966) gave all subjects a second recall test which was always a cued recall test. They found, of course, that cued recall showed a large increase over a previous noncued recall for the usual reason—the cues gave access to additional categories. Of greater interest are comparisons among the three cued tests: test 1 cued, and test 2 cued after either uncued or cued test 1. Tulving and Pearlstone found test 2 cued after test 1 cued gave recall practically identical to test 1 cued. However a previous *uncued* test seemed to reduce test 2 cued performance. This is illustrated in Fig. 20 at the *top*. The differences were not explained by Tulving and Pearlstone.

In the bottom of Fig. 20 we show the predictions when the present SAM model is applied to the second test. In this application the final test matrix after the first test (uncued) was for each simulation run the starting

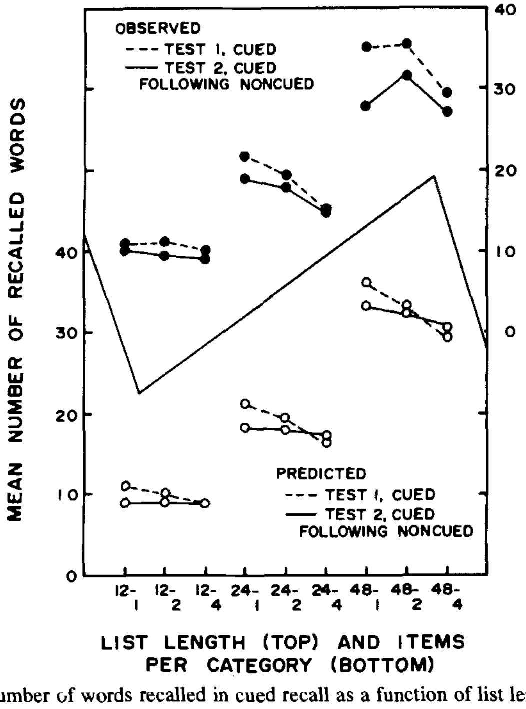


Fig. 20. Mean number of words recalled in cued recall as a function of list length and words per category for subjects who were given these cues either on the first recall test, or not until the second recall test. (Data from Tulving & Pearlstone, 1966.) Predictions from simplified SAM model, described in text.

test matrix for the second test (cued). Otherwise the second test assumptions were identical to those used for a cued first test (including the occurrence of new incrementing for items recalled in test 2, regardless of whether test 1 incrementing had already occurred). Clearly the predictions are matching the main features of the data, but why does a noncued first test hurt a cued second test? The answer depends upon the context-item increments that take place in test 1. In test 2 these increments tend to cause sampling of the previously recalled items in test 1, to the exclusion of items (and categories) that were not recalled in test 1, especially when category size is small.

There is one other set of important findings in categorized free recall that may be considered in the context of our model. These findings concern cued recall; they show that the probability of recall of a category member decreases slightly but systematically as successive categories are cued. This result seems to have been found first (independently of each other) by Dong (1972) and Smith, D'Agostino, and Reid (1970). More systematic studies are reported in Smith (1971) and Roediger (1973). In the experiments of Smith (1971) blocked presentation of categorized words was followed by cued recall. A significant decline in word recall for successive categories tested was observed. This output interference effect was not dependent on the inclusion of the last input category nor was it decreased by introducing an interpolated task between study and test. Thus the results cannot be attributed to a short-term forgetting process. More output interference was observed when a long recall time per category (60 sec) than when a short time (30 sec) was given. In one of Smith's experiments (Smith, 1971, Exp. IV) presentation time per item and category size were varied in a between-list design. More output interference was observed with longer categories and with a higher presentation rate. Roediger (1973) varied category size within a single list and found no effect of category size. Thus, one may conclude that the effect depends on the absolute number of items previously recalled rather than on the number of items per category stored in memory.

Roediger (1973) observed that the probability of recall for successively tested categories decreased in an approximately linear fashion with a slope of about −.007. Figure 21 shows the predictions of our model for cued recall for a list of 20 categories of 4 items each. These results are based on 1000 simulation runs with the parameter values that were estimated from the results of Tulving and Pearlstone (1966). It is evident that the model predicts this output interference effect. In our simulation we obtained a slope of −.0063, so the magnitude of the effect is also predicted quite well.

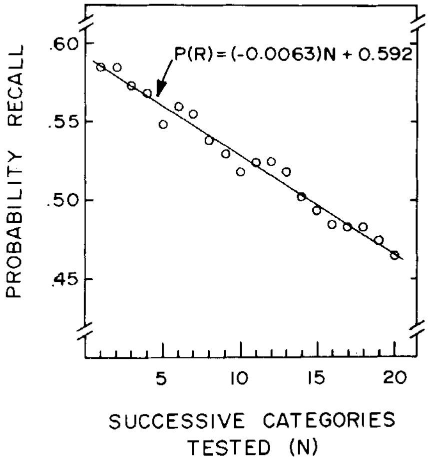


Fig. 21. Predicted probability of recall as a function of the output (i.e., test) position of the category in cued recall. Best fitting regression line is shown. The model assumed 20 categories of 4 words each, and used parameter values equal to those utilized for the fit to the Tulving and Pearlstone (1966) data (see Figs. 17–20, and the text).

The output interference prediction is due to the *incrementing* of context-to-item associations, when there are residual associations between category cue and nc-image. Later in the sequence of tested categories, there is an increasing tendency to sample recalled items from earlier categories, because their associations to context had been incremented. The model also predicts that the effect will be stronger the lower the initial strength of the context associations because the increment after retrieval will than be relatively higher. Thus, the model predicts that higher presentation rates (i.e., shorter presentation times per item) should lead to more output interference as was found by Smith (1971). For these same reasons, our model is consistent with the result obtained by Roediger (1973) that the output interference effect depends on the absolute number of items recalled previously.

A similar explanation handles an effect noted by Roediger (1978): providing some of the category names as retrieval cues increased the number of words recalled from the cued categories [i.e., the positive cuing effect observed by Tulving & Pearlstone (1966)], but decreased the number of words recalled from the noncued categories (i.e., the negative cuing effect observed by Slamecka and others, see below). In this case, there will be a tendency to sample the words recalled earlier from the cued categories, due to incrementing. (Note that such an explanation does *not* handle Slamecka's part list cuing effect, however, as discussed in the next section.) Verification of our reasoning comes from another condition in Roediger (1978): one group was given the category names with the instruction not to recall from those categories and another group was given those names with the instruction to recall especially from those categories. Relative to a control group a large decrement in the number of critical words recalled was observed for the second group but not for the first group. These results show that it is the act of recall that produces the interference.

If the category cue to nc-image residual, *D*, is raised, the model predicts cued performance to drop, but the magnitude of the output interference effect to remain virtually unchanged. For example, when we doubled the value of *D* in the simulation, the recall level in Fig. 21 dropped by 10%, but the slope of the function did not change. One way to increase *D* experimentally is to use categories that are more similar to one another. Roediger and Schmidt (1980, Exp. III) carried out such a study, and found just this predicted pattern of results. (Roediger & Schmidt, 1980, Exp. IV, showed a similar effect in cued recall of paired associates, a finding matching the predictions of the theory, as shall be described in the section on paired associates below.)

To summarize all these findings concerning categorized free recall, our very simple SAM model, without interitem associations, proved capable of predicting all the major results from this paradigm, including cued, and uncued, and partially cued recall, and the output interference effect. Furthermore, these results and simulations demonstrate clearly the need for residual associations between category cues and noncategory items from the list, to explain list length effects on within-category recall, and to explain the output interference effect. In addition, the need for incrementing is clear, to explain the output interference effect. Thus, although these factors are not needed in the model to deal with simple free recall of uncategorized lists, both residual associations and incrementing are necessary components of SAM. (The basis for interitem residuals will be discussed later.)

Finally, one might ask whether additions to our simplified model, such as word cues and interitem search routes, or recalls from categories different from the cue, can present any difficulties for SAM. Such changes add quite a few processes and parameters to the model; we have applied such an extended model to this data, with equal success, but do not present the results since no new insights are gained. However, we shall give such an extended version of the model when turning to our next application, where the two categories in a list consist of (1) random words, and (2) complex pictures.

D. FREE RECALL OF PICTURES AND WORDS

The study on pictures and words was carried out and analyzed by Gary Gillund at Indiana University. We give only a brief resume of the major points here (see Gillund & Shiffrin, in preparation). Some lists contained only words, others only pictures, and some contained some of both. The numbers of words and pictures in these various lists were covaried. In mixed lists, presentation was either blocked, or was alternated as evenly as possible. Arithmetic was used after presentation and before recall to empty STS. Recall of pictures was obtained by the method of Shiffrin (1973): the subjects wrote very brief descriptions of each recalled picture. Then, after recall was completed, all the list pictures were shown to the subjects, who matched their descriptions to the pictures. List lengths used were 10 and 20, and presentation time was 2 sec per item.

Some of the main results are shown in Figs. 22 and 23. Note that the usual list length effects were obtained in both pure and mixed lists, but

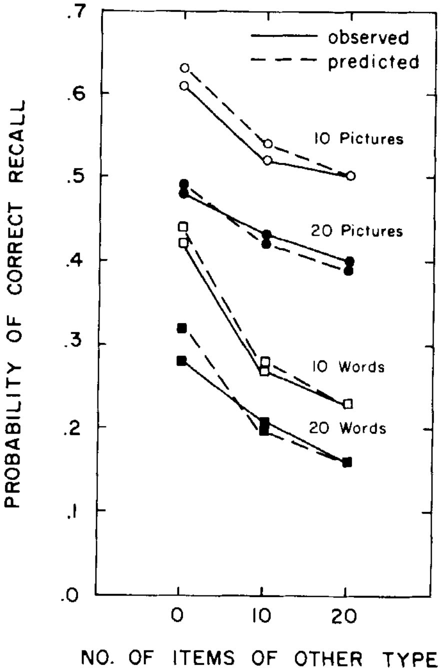


Fig. 22. Probability of recall for pictures and words as a function of both category size and other-category size. Predictions from the SAM model described in text.

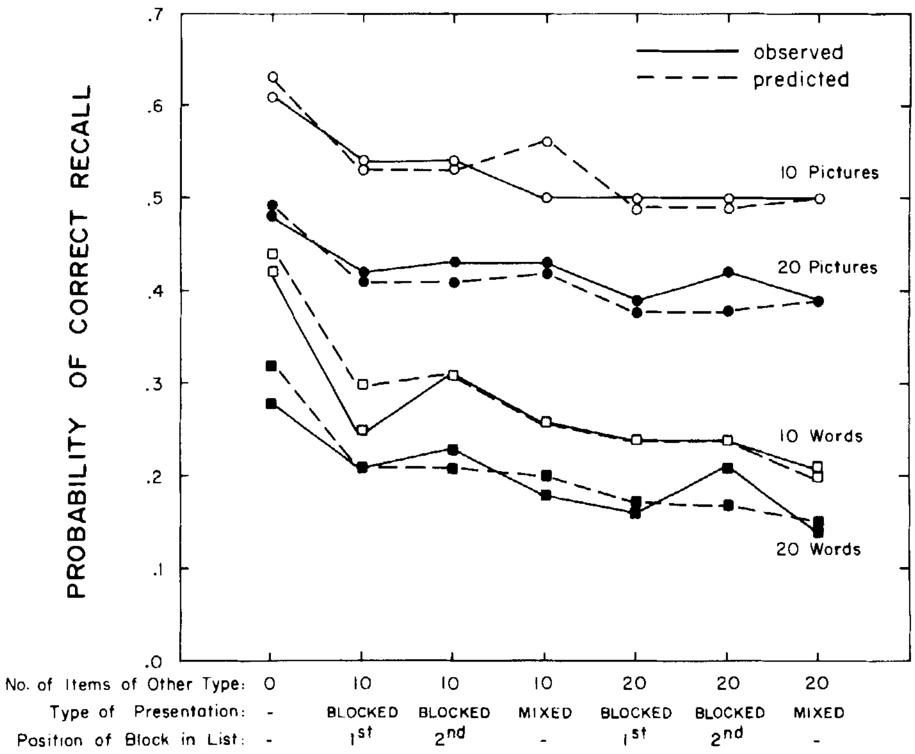


Fig. 23. Probability of recall for pictures and words, as a function of blocking order of testing of each block, and mixing, for each category size, and other-category size. Predictions from SAM model described in text.

also a large effect was obtained for recall of pictures when only the number of words was varied, and vice versa. This is of some interest since the words were chosen to be of low imagery value, and the pictures were complex and not easy to describe succinctly or accurately in words.

The model applied to this case is the basic SAM model for two categories (pictures and words), but with all residual associations, inter-item searches, category searches, and so forth, included. The storage process utilizes a buffer of size 4 for words, and size 1 for pictures; whenever a picture is presented, the previous buffer contents are cleared; when a picture is in the buffer a presented word replaces it. Words build up item-context (a_w) and item-item (b_w) strengths as a function of rehearsal time; pictures build up item-context strength (a_p) as a function of rehearsal time. An item's self-association strength is set equal to the context strength. Residual associations fill the rest of the test matrix: picture-picture (d_{pp}), picture-word (d_{pw}), word-picture (d_{wp}), and word–word (d_{ww}, for jointly unrehearsed words). In addition, it is assumed that a category cue may be used during retrieval. Thus a within category residual, d_{cc}, and a residual from a category cue to an item in the other category, d_{nc}, are also assumed.

The retrieval plan is fairly straightforward. When a new item is recalled, the next cue-set consists of the recalled item, its category, and context. The search begins with the context cue and a randomly chosen category cue. Any time a cue-set including a word category cue fails L_{MAXW} consecutive times, or a cue set including a picture category cue fails L_{MAXP} consecutive times, the next cue-set consists of a change in category cue, and context (no item cue).

When an item is recalled, increments of the cue to image associations take place: context–word e_w; context–picture e_p; word–word and word–category–word, both f_{ww}; picture–picture and picture–category–picture, both f_{pp}; word–picture and word–category–picture, both f_{wp}; picture–word and picture–category–word, both f_{pw} (self-increments were equated to context increments). When the total failures (including those accumulated during item cue searches) reach K_{MAX}, search ceases. No rechecking is assumed.

One could guess that this model has the power to deal with the data in Figs. 22 and 23. In fact, many combinations of parameters give more or less equivalent fits to the data. Just as with the category data, more insights into the model would probably be gained by applying simpler versions of the model. Nevertheless, the assumed processes and parameters represent logically necessary components of the SAM model, and it is of some use to show the predictions of the complete model. The predictions in Figs. 22 and 23 represent the results of a rough parameter search, with $a_w = .29$; $a_p = .56$; $b_w = .01$; $d_{pp} = .0305$; $d_{wp} = .03$; $d_{ww} = .02$; $d_{pw} = .025$; $d_{cc} = .0385$; $d_{nc} = .005$; $e_w = .03$; $e_p = .007$; $f_{ww} = .03$; $f_{pp} = .007$; $f_{pw} = f_{wp} = .001$; $L_{MAXW} = 3$; $L_{MAXP} = 4$; $K_{MAX} = 32$. Quite clearly the model captures the main aspects of the data. It is reasonable to conclude that complex pictures, and words, are comparable entities that may be treated similarly in retrieval (at least within the context of a SAM-like model).

E. PART-LIST CUING

One of the more remarkable findings in free recall, primarily because it does not seem consistent with traditional associative theories of memory, is known as the part-list cuing effect (e.g., Slamecka, 1968, 1969). Suppose after list presentation that a random subset of the list items is presented to the subjects in the cued group, who are told to use them as cues to aid in recall of the remaining list-items, called "critical items." The control group is given no cues, and recalls freely, as usual. The control group actually recalls slightly *more* critical items than the cued group. This finding is robust, being found consistently in random lists, and within categories of categorized lists.

The mystery of the control group advantage was first discussed by Slamecka (1968, 1969). He argued that at least some of the critical items that were not recalled by the control group should have been recalled by the cued group due to the presence of cues that would not have been retrieved by the control group. This argument depends upon the formation during storage and use in retrieval of interitem associations. Slamecka and many later theorists therefore concluded from the part-list cuing effect that such interitem associations could not have been both stored and used in retrieval.

This reasoning is not, however, correct. We shall show next that a prediction of the part-list cuing effect is inherent in SAM-like models, despite the heavy use of interitem associative structure that is made in such models. In fact, it is this very structure and its use in retrieval that produces the effect. This entire problem is discussed in all possible variations, and the literature thoroughly reviewed, in Raaijmakers and Shiffrin (1981). In this article, therefore, we shall summarize these matters in very brief fashion.

One of the most surprising findings related to the part-list cuing effect concerns the effect of increasing the similarity of the list items to each other. Slamecka (1968, Exp. VI) used three lists: (1) 30 rare words; (2) 30 common words; (3) a list consisting of "butterfly" and 29 of its most popular associates. The control groups recalled 5.58, 7.04, and 8.50 critical words, respectively. The cue groups recalled 4.70, 6.79, and 8.97 critical words, respectively. Thus increasing similarity almost doubled recall, while only slightly altering the basic effect.

The SAM model will now be fit to an idealized part-list cuing paradigm. We assume that 30 words are presented (of varying interitem similarity in different lists), followed by arithmetic. The control group free recalls as usual. The cued group is given 15 randomly chosen words from the list, and told to use them to aid recall of the remaining words.

The model for the control condition is identical to that used for Robert's data discussed earlier; even the parameters are identical, except that the interitem strength parameter, *b*, is systematically varied for different lists. The model for the cued condition is almost the same, except that the subject is assumed to use the provided cues before reverting to the normal search. In particular, each provided cue is used in the cue-set, along with context, until L_{MAX} failures are reached. A recovery of another cue word is *not* counted as a failure, except on the second and subsequent recoveries of the same cue word (to make the cue condition comparable to the control condition). Recalled items during this phase of cued search are simply "written down" but not used as cues themselves. If K_{MAX} is not reached when the provided cues are used up, then normal search commences, as in the control condition. When K_{MAX} failures are reached, then all previously recovered items (whether cues or critical items) are rechecked. All parameters are the same as in the control condition.

The predictions for the cued and control conditions, for various values of the interitem strength parameter, *b*, are shown in Fig. 24. Note that recall is predicted to double as *b* increases, but the control group advantage decreases only slightly over the same range.

In Raaijmakers and Shiffrin (1981), these predictions are exhaustively explored, through numerous versions of the SAM model. The basic predictions hold without incrementing, without rechecking, regardless of the particular rechecking assumptions made, whatever the value of K_{MAX} or L_{MAX} (within reason) and regardless of the particular stopping rule used, among other variations. Surprisingly, the control group advantage occurs in the face of a factor favoring the cue group: when an item-cue plus context cue are used, recovery probability is higher than when a context

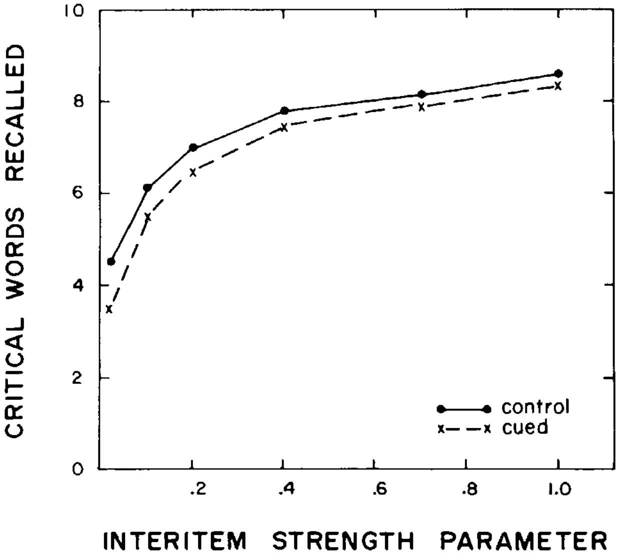


Fig. 24. Predictions for the part-list cuing effect for an idealized paradigm, as interitem strength parameter, *b*, is varies. List length is 30, number of cues is 15, presentation time per word is 2 sec; *d* is set to equal .2*b*. Other parameters as in Fig. 8.

cue is used alone; since the cue group uses more item searches, it obtains an advantage. For example, if recovery probability is set equal to .75 regardless of the cues used, then the cued group inferiority increases by about one-half a word.

What then explains SAM's prediction of the part-list cuing effect? The main factor is illustrated by Fig. 25. Suppose that LTS after presentation of 12 words consists of 4 triads, unrelated to each other but so strongly interrelated that recall of any one item in a triad leads to recall of the entire triad. The critical items and cues are indicated in the figure. Suppose that during search of this structure, the control and cued groups sample an equal number of triads (a simplification for the sake of the argument). The cued group's sampled triads will all contain a minimum of one cue word, and hence a relatively small number of critical words. The control group's sampled triads, on the other hand, will often contain no cue words and hence be relatively rich in critical words. Note well that

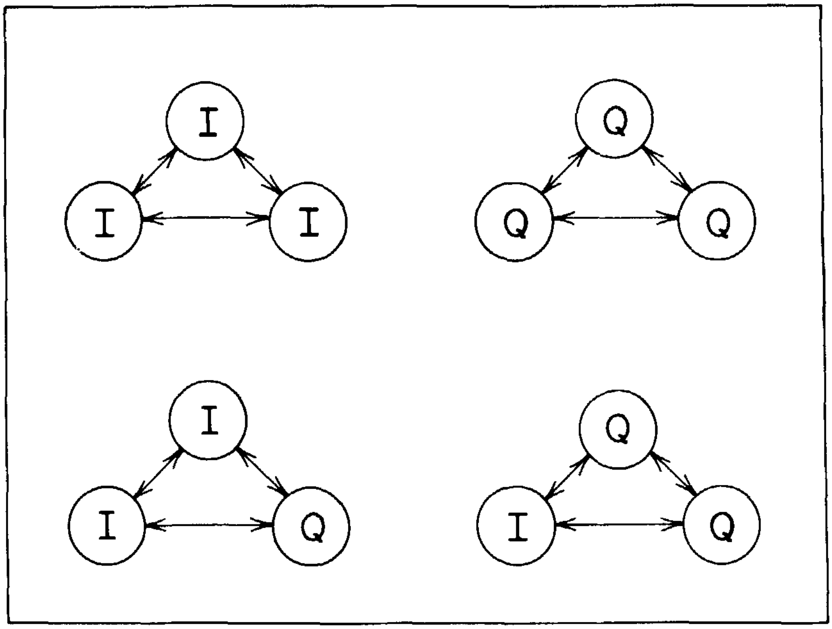


Fig. 25. A simplified associative network for a 12-word list stored as four triads. The six experimenter provided cue words have images denoted Q. The six remaining critical items are denoted I. The arrows denote associations between images. Each image has an association to context, not depicted. A context sample can access a triad rich in critical items (e.g., the triple-I triad). The cue word plus context samples can only sample triads relatively impoverished in critical items, since each such triad must contain at least one cue.

this effect depends upon both a nonuniform structure in LTS, and also extensive and effective use of interitem search.

SAM's prediction of the part-list cuing effect is thus dependent upon the fact that both groups make extensive use of interitem search. The control group uses a mixture of two types of cue sets; one type consists of context cues only, while the other type consists of both a context cue and a word cue that was generated by the subject earlier in the search. On the other hand, the cue group usually uses just one type of cue set, containing both a context cue and a word cue provided by the experimenter. The control group is superior under these circumstances for the reasons given above. Strangely, then, the control group advantage is the result of just the interitem structure and interitem retrieval routes that previous theorists have argued must be ruled out.

We mention, finally, that several other findings in part-list cuing are predicted by SAM. It has been found (especially in categorized lists) that increasing the number of provided cues slightly increases the control group advantage (e.g., Slamecka, 1968; Roediger, Stellon, & Tulving, 1977; Roediger, 1974). To apply SAM, we reduced L_{MAX} to 2 and raised K_{MAX} to 50, to ensure that all provided cues would always be used, and studied the predictions as the number of cues from a 30 item list varied from 0 to 25. The predictions (all other parameters as in Fig. 24) are given in Fig. 26. Clearly the model predicts a slight, almost linear, decrease as the number of cues increases.

A version of SAM has also been applied to the categorized list paradigm, in which cues may be provided from each category. Furthermore, these cues may be "extra list," from the relevant category, but not on the list. Bruce Williams at Indiana has applied the model to this situation. The model is similar to, and an extension of, the model fit by Gary Gillund in Section III,D to the two category situation. It is described in Raaijmakers and Shiffrin (1981). We show here the results only when the model was fit to data of Watkins (1975, Exp. 1). Figure 27 gives the data and predictions. Clearly the effects of number of cues and extra list cues are both well predicted by SAM.

One effect found by Mueller and Watkins (1977) could not be predicted by SAM, despite variations in parameters and assumptions. This effect was the fact that cues from categories other than the tested category did not produce a disadvantage in comparison with the control condition. For assumptions and parameter values which *would* produce a cue condition disadvantage for within-category, within-list, cues and for within-category, extra-list cues, a disadvantage was also predicted for the extra category within-list cues. Thus the fact that a particular version of SAM contains a profusion of parameters and processes does not necessarily

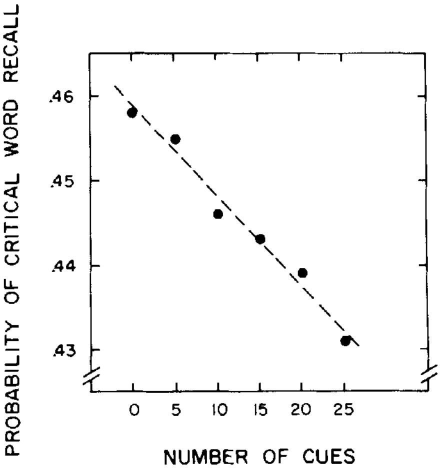


Fig. 26. The same model as in Fig. 24, with $b = .10$, $L_{\mathrm{MAX}} = 2$, $K_{\mathrm{MAX}} = 50$. The probability of critical word recall is given as a function of the number of experimenter-provided cues (list length = 30).

reduce the testability of the model. Why does SAM mispredict the extra-category cuing result? We suggest that subjects realize that the cues are from a different category, and thinking such cues to be worthless, ignore them. We suggest that a disadvantage would appear if only the subjects could somehow be induced to *use* the provided cues.

F. PAIRED-ASSOCIATE PARADIGMS

The reader will undoubtedly have noticed that the model developed for free recall contains all the ingredients necessary to predict cued testing of paired associates. We have in fact embarked upon an extensive research program in which paired associates are presented and tested by various methods (free recall, cued recall, recognition). In this article we will present only the first of these studies, and that in brief fashion (see Raaijmakers, 1979, for additional details).

The study was a natural generalization of the simple free recall paradigms that have been discussed. Pairs of items (paired-associates, denoted PA) were included in lists along with single items (denoted FR, for consistency with the previous studies). A single trial procedure was used, so each word was seen only one time. The 10 conditions were:

Number of PA items:	5	5	15	15	5	15	30	0	0	0
Number of FR items:	10	30	10	30	0	0	0	10	30	40

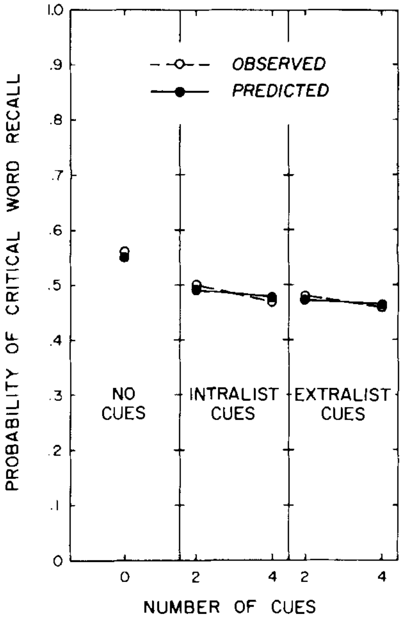


Fig. 27. Predicted and observed probabilities of critical word recall in a category, for the control condition, and for the intralist and extralist cue conditions, each with two or four cues. (Data from Watkins, 1975.) The model is described in Raaijmakers and Shiffrin, 1981; parameters: list length = 36; words per category = 6; presentation time per word = 3 sec; $r = 4$; K_{MAX}/category = 12; $L_{\mathrm{MAX}} = 3$; $a = c = .38$; $b = .38$; item-category strength per sec = .38; category-cue and list-word-cue residuals to words on list = .1; all increments = .36; residual strength of extralist cues to list items in same category = .03; product of residual strengths when self-sampling an extra-list cue = 2.2.

(Note that the number of PA items is given in terms of the number of pairs; the number of words is therefore given by twice this number.) PA pairs and FR items were randomly mixed. For each condition, half the subjects were first given cued testing of the PA items (paired-associate testing), followed by free recall of the FR items; the other subjects were tested first on the FR items (free recall testing), followed by cued testing of the PA items. Subjects were not told before study of a list which items would be tested first. The words were presented visually, a single word for 2 sec, a pair for 4 sec. Paired words were tested either in a forward manner or in a backward manner: if the pair was A-B it was tested either as A-? or as ?-B.

Subjects were asked to allot an equal amount of effort in studying each word. The instructions emphasized that they should try to link together the two members of a word pair into a single unit, by forming a mental image or by using some kind of verbal code. After presentation of the list a 20-sec arithmetic task was given to eliminate short-term effects. A written recall procedure was used. Single words were tested using a 2-min free recall procedure, paired words were tested with a paired-associate testing procedure. In this case the subjects had 4 sec to write down their answer.

Figure 28 in the top panel shows the effect of list length on recall of the PA items and the FR items. These data are averaged over order of testing and over testing with the A member and with the B member of the A–B pair. It is evident that the results are quite consistent. In free recall testing the probability of recall decreases not only as a function of the number of FR items but also as a function of the number of PA items on the list. A similar list-length effect is observed for the PA items, and again the probability of recall decreases when other items are mixed in the list. In contrast with the word–picture study, where recall was not directed to one category or the other, the present "cross-category" list length effects take place even though recall is directed specifically to either FR or PA items. The results are similar to those found in cued recall of categories, where recall depends upon the number of items in the other categories.

Figure 29 gives the effects of test order. The main result to note here is

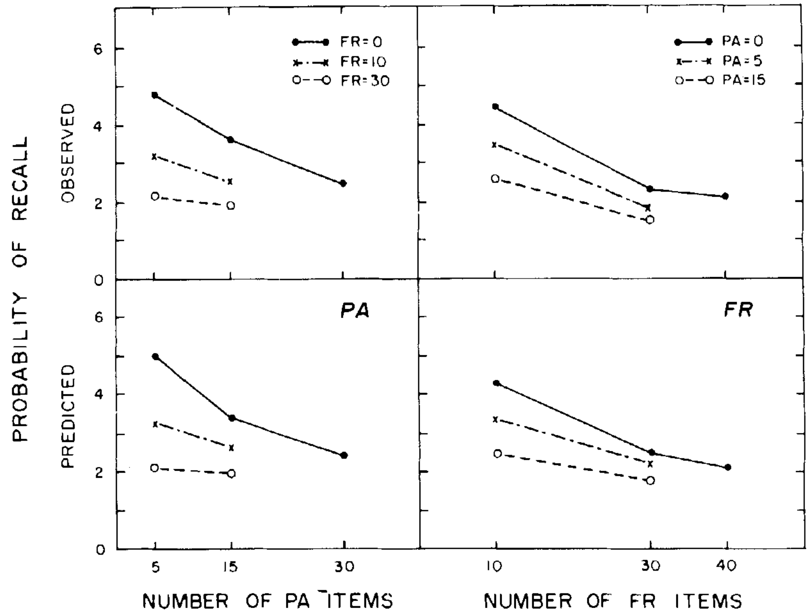


Fig. 28. The effect of the number of single words (FR) and the number of words pairs (PA) on the probability of recalling these two types of items. Predictions based on the SAM model described in the text.

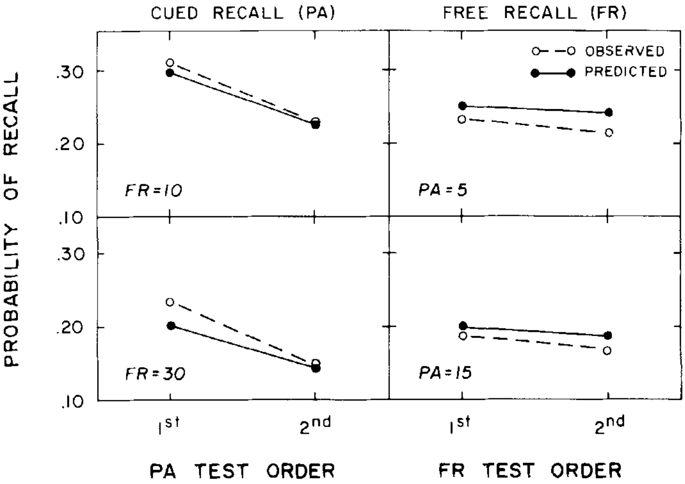


Fig. 29. The effect of the order of the FR and PA tests upon the probability of recall of each type, partitioned according to the number of items of the *other* type. Predictions based on the SAM model described in the text.

that free recall of the FR items is only slightly reduced by a prior cued test of PA items (about .02 reduction in probability of recall, overall). On the other hand, cued recall of PA items is considerably reduced by a prior free recall test of FR items (about .08 reduction in probability of recall, overall).

The final results to which we wish to call attention are given in Fig. 30; averaged across all conditions, this figure gives the overall probability of cued recall of PA items, broken down by test quartiles. That is, this figure shows that cued recall probability drops slightly as the test position of the pair is delayed.

Applying SAM to this study is quite easy, since all the groundwork has already been laid in the models for free recall. The storage assumptions are straightforward. Each pair of PA items clears the buffer. Each FR item clears the buffer of PA items, but adds to any FR items already in the buffer (up to the buffer size, r: then one of the previous buffer members is deleted). A PA pair builds up item-context strength (parameter = a_{PA}) and interpair-strength (parameter = b_{PA}) as a function of rehearsal time (always 4 sec in this case). An FR item builds up item-context strength (parameter = a_{FR}) and item-item strength (parameter = b_{FR}) as a function, respectively, of rehearsal time, and of joint rehearsal time. Note in each case that if there are m individual words in the buffer together for t sec, the rehearsal time for any word, or any pair of words, is equal to t/m.

The remainder of the test matrix is filled with various residual associa-

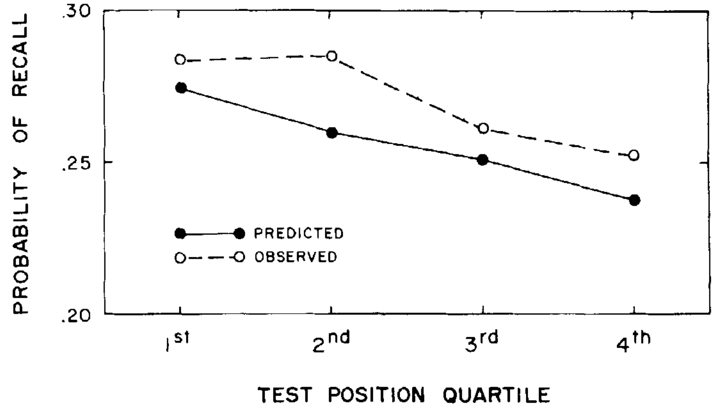


Fig. 30. Predicted and observed recall probability of PA items, for all list types lumped together, as a function of test quartile. Predictions based on model described in text.

tions. Any item's self-association strength was set equal to its context strength. To keep things simple, *all* strengths not mentioned above were set equal to a common residual value, *d*.

Finally, various increments are assumed after a successful recall. Recall of an FR item is followed by increments of e_{FR} for the context and self-strengths, and f_{FR} for the interitem strength if an item cue is used in the cue-set. Recall of a PA item is followed by an increment of e_{PA} for the context and self-strengths (and presumably by an increment of f_{PA} for the interitem strength, but this parameter is never needed in this application—see below).

Since intrusions of PA items in free recall, or FR items in cued recall, were very rare, we assume that any sampled and recovered item can be classified correctly as to type. Hence any sample of a "wrong" type during search is counted as a failure (and no incrementing occurs). With this proviso, the retrieval model for free recall is identical to that described for, say, Robert's data. For cued recall, it is assumed that each sample is made with both context and the provided item cue, and search ceases when L_{PAMAX} failures is reached (since only 4 sec were provided by cued recall, L_{PAMAX} was arbitrarily set to 1.0). Note that any increment between a PA cue and the correct response has no effect within this model, because such a cue will never be used again, either in PA or FR testing.

Only a very limited parameter search was carried out, but a reasonably adequate fit was obtained for the following values: $K_{MAX} = 30$; $L_{MAX} = 3$; $L_{PAMAX} = 1$; $r = 4$; $a_{FR} = .30$; $a_{PA} = .18$; $b_{FR} = .30$; $b_{PA} = .60$; $d = .035$; $e_{FR} = 3.0$; $e_{PA} = .80$; $f_{FR} = 3.0$. The predictions are shown in

Figs. 28, 29, and 30. It seems clear that the SAM model, essentially the same SAM model used for free recall, can handle the major findings of cued PA recall, even for the case when FR items and PA pairs are mixed within the same list. The list-length effects, the effects of order of PA and FR testing, and the effects of test order in PA testing are all well handled by the model.

In many ways, the theoretical conclusions to be drawn from this study and the application of the model parallel the conclusions reached from the categorized free recall situation. The list-length effects that cross test type, and that appear in cued recall, illustrate the importance of residual associations between items not rehearsed together, and even between items of different types. The effects of order, in both Figs. 29 and 30, illustrate the importance of the incrementing process. Of course, the basic phenomena of cued testing require interitem associations and search routes. Finally, we regard it as a strong point in favor of the model that a system developed for free recall can handle so accurately these various results from cued testing of paired-associates.

IV. General Discussion and Final Comments

We begin by calling attention to a problem that Smith (1978) has termed "the sufficiency/transparency tradeoff." The problem is that as a long-term memory model (especially a simulation model) becomes more and more complex, and increasingly encrusted with special assumptions, it gains the ability to predict a good deal of data (sufficiency), but becomes increasingly opaque to external observers (including the model's creators). That is, it becomes virtually impossible to extract the essential principles from the mass of details and interactions that comprise the model, and it is often impossible to anticipate what the model will predict for a given manipulation.

We have been quite concerned, even for our relatively simple model, with the "transparency" problem, and have adopted a series of measures to deal with it. First, we do not attach much significance to the fact that the model can fit any single study or type of study. Rather, we require the model to apply to many different tasks and types of tasks, with essentially the same set of assumptions, and the same set of basic mechanisms. Second, if the values of parameters are important to predict certain effects, those values should be consistent with the model's rationale and the task requirements. Third, the model should have testable aspects—there should be some results that the model cannot fit (the part-list cuing effect is an example of an inherent prediction of our model—in fact, we saw that the failure to attain a cue condition disadvantage for extracategory cues could not be handled by the model). Fourth, and perhaps most important, the model must be made understandable to the observers. We have attempted to do this by carrying out extensive explorations of the "assumption space" of the model, systematically adding and removing various processes, and examining the shifts in predictions that result. Our model, unfortunately, is so stochastically interactive, that it is difficult to make accurate intuitive predictions even for very simple combinations of assumptions. We have seen this to be the case especially when applying the model to "hypermnesia" and "part-list cuing." In light of our theoretical explorations, we hope the predictions of these effects, and indeed the basic workings of the model, have been illuminated.

Let us review now the basic tasks to which the model has been applied. Serial position effects, but more important, list-length and presentation time effects in single-trial free recall were easily handled. The temporal aspects of free recall were dealt with next, including the effects of instructions to extend the period of active retrieval, cumulative response curves, repeated recall, the effect known as "hypermnesia," and interresponse times. SAM was applied next to the basic phenomena in categorized free recall, not only handling the large effects of cuing, of category size, or number of categories, of mixtures of pictures and words, of the number of categories upon within category recall, and of the test order of categories (the output interference finding), but also explaining the subtle effects of cued recall following noncued recall. The model was next shown to predict the part-list cuing effect in its sundry variations, an important result since previous associative models have had difficulty dealing with the finding. Furthermore, the explanation was not post hoc; the model for free recall was applied "intact" to the part-list cuing paradigm, and the prediction proved to be an inherent property of the model, occurring in almost all model variations. Finally, the model was shown to predict cued recall of paired associates, in lists containing both paired-associates and single items. Since the model for free recall utilized extensive amounts of item-cuing, the extension to the paired-associate situation required no new assumptions. The predicted effects include those of list length, number of PA items, number of FR items, sequential effects during cued testing, and the relationship of free recall to cued recall for different test orders.

These are not the only tasks to which the model has been applied, but space restrictions prevent our presentation of these other paradigms. In brief, they include a variety of other paired associate tasks, and several recognition paradigms. Recognition may well involve an initial judgment of "familiarity," perhaps based on the value of the denominator of the sampling equation. If familiarity does not lead to a response, however, then the rest of the search is treated similarly to that for recall.

Let us conclude by reprising the important features of our retrieval theory. An associative retrieval structure and cue-dependent retrieval are essential, but are common to many theories. The sampling assumptions are the key to the present approach, in several different ways. First, the fact that sampling is probabilistic allows for a considerable degree of resampling in certain circumstances. Such resampling of previously sampled images is the basis for stopping the search, and hence an important contributor to the limitations upon retrieval. Second, the sampling equation [Eq. (1)] provides an explicit basis for combining cues. That is, the multiplication of strengths in an additive ratio rule provides a means of focusing the search when necessary or desired, and allows SAM to predict cued or free recall with equal facility. Turning now to recovery, it is obvious on logical and empirical grounds that some type of recovery rule is necessary (for example, the effects on free recall of doubling presentation time would be most difficult to handle without a strength-dependent recovery probability).

These factors notwithstanding, we make no claim for the uniqueness of the particular mathematical forms of Eqs. (1) and (2). These functions were chosen for simplicity, convenience, and historical factors, but slight variations in their forms would undoubtedly lead to an equally good description of most of the data. It is our position that the basic framework of the model has enough power to handle the data that small variations in quantification will do little to degrade the quality of the predictions. A test of this position must await further empirical and theoretical work.

Let us turn now to some of the subsidiary assumptions of SAM. The inclusion of residual associations makes our retrieval network "completely" interconnected, a rather novel feature. Such interconnectivity is needed to explain list-length effects in various types of cued recall, in both the category and paired-associate paradigms. Incrementing represents learning effects that occur during retrieval; it is essential to explain various types of test-order findings (as in successive testing of categories, for example). Still other factors in our model do not seem crucial for predicting present data, or have not yet been explored theoretically. Such factors include the conditionalization rules that apply after resampling of the same image and rechecking.

Finally, there are subject controlled strategies in our theory, such as search termination rules, and choice of cues at various stages of the search. We have tried to include reasonable strategies in our simulation, but are convinced that a theory would be very weak if its predictions depended in important ways on the choice of particular strategies (since different subjects probably choose different strategies, and the same subject probably changes strategies from time to time). It is for this reason that we have expended considerable effort in this article showing the effects upon the predictions of changing strategies. It is one of the successes of this model that the basic predictions are quite insensitive to "sensible" alterations in retrieval strategies, but that manipulations expressly designed to change strategies (such as encouragement to continue search) have effects in the data that are predicted through simple manipulations of the appropriate parameter in the model.

ACKNOWLEDGMENTS

This research was supported in part by a fellowship from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.) to the first author, and Public Health Service Grant 12717 and National Science Foundation Grant BNS-77-00156 to the second author. We are indebted to Gary Gillund and Bruce Williams who carried out part of the research described in this article.

REFERENCES

Anderson, J. R. FRAN: A simulation model of free recall. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 5). New York: Academic Press, 1972. Pp. 315–378.

Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D. C.: Winston, 1973.

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press, 1968. Pp. 89–195.

Bousfield, W. A., Sedgewick, C. H. W., & Cohen, B. H. Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology*, 1954, **67**, 111–118.

Bower, G. H. Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D. C.: Winston, 1972. Pp. 85–123.

Cohen, B. H. Recall of categorized word lists. *Journal of Experimental Psychology*, 1963, **66**, 227–234.

Cohen, B. H. Some-or-none characteristics of coding behavior. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 182–187.

Donaldson, W. Output effects in multitrial free recall, *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 577–585.

Dong, T. Cued partial recall of categorized words. *Journal of Experimental Psychology*, 1972, **93**, 123–129.

Erdelyi, M. H., & Kleinbard, J. Has Ebbinghaus decayed with time? The growth of recall (hypermnesia) over days. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, **4**, 275–289.

Estes, W. K. Statistical theory of spontaneous recovery and regression. *Psychological Review*, 1955, **62**, 145–154.

Gillund, G., & Shiffrin, R. M. Free recall of pictures and words. In preparation.

Indow, T., & Togano, K: On retrieving sequence from long-term memory. *Psychological Review*, 1970, **77**, 317–331.

Johnson, D. M., Johnson, R. C., & Mark, A. L. A mathematical analysis of verbal fluency. *Journal of General Psychology,* 1951, **44,** 121–128.

Luce, R. D. *Individual choice behavior: A theoretical analysis.* New York: Wiley, 1959.

Mueller, C. W., & Watkins, M. J. Inhibition from part-set cueing: A cue-overload interpretation. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 699–709.

Murdock, B. B., Jr. The serial position effect in free recall. *Journal of Experimental Psychology,* 1962, **64,** 482-488.

Murdock, B. B., Jr., & Okada, R. Interresponse times in single-trial free recall. *Journal of Experimental Psychology,* 1970, **86,** 263-267.

Raaijmakers, J. G. W. *Retrieval from long-term store: A general theory and mathematical models.* Doctoral dissertation, University of Nijmegen, Nijmegen, The Netherlands, 1979.

Raaijmakers, J. G. W., & Shiffrin, R. M. Search of associative memory. *Psychological Review,* 1981, in press.

Roberts, W. A. Free recall of word lists varying in length and rate of presentation: A test of total-time hypotheses. *Journal of Experimental Psychology,* 1972, **92,** 365–372.

Roediger, H. L.,III. Inhibition in recall from cueing with recall targets. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 644–657.

Roediger, H. L.,III. Inhibiting effects of recall. *Memory and Cognition,* 1972, **2,** 261–269.

Roediger, H. L., III. Recall as a self-limiting process. *Memory and Cognition,* 1978, **6,** 54–63.

Roediger, H. L., III, & Schmidt, S. R. Output interference in the recall of categorized and paired-associate lists. *Journal of Experimental Psychology: Human Learning and Memory,* 1980, **6,** 91–105.

Roediger, H. L.,III, Stellon, C. C., & Tulving, E. Inhibition from part-list cues and rate of recall. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, **3,** 174–188.

Roediger, H. L.,III, & Thorpe, L. A. The role of recall time in producing hypermnesia. *Memory and Cognition,* 1978, **6,** 296–305.

Rundus, D. Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 43–50.

Shiffrin, R. M. Memory search. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970. Pp. 375–447.

Shiffrin, R. M. Visual free recall. *Science,* 1973, **180,** 980–982.

Shiffrin, R. M. Short-term store: The basis for a memory system. In F. Restle, R. M. Shiffrin, N. J. Castellan, H. Lindman, & D. B. Pisoni (Eds.), *Cognitive theory* (Vol. 1). Hillsdale, NJ: Erlbaum, 1975. Pp. 193–218.

Shiffrin, R. M. Capacity limitations in information processing, attention and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 4): *Memory processes.* Hillsdale, NJ: Erlbaum, 1976. Pp. 177–236.

Slamecka, N. J. An examination of trace storage in free recall. *Journal of Experimental Psychology,* 1968, **76,** 504-513.

Slamecka, N. J. Testing for associative storage in multitrial free recall. *Journal of Experimental Psychology,* 1969, **81,** 557–560.

Smith, A. D. Output interference and organized recall from long-term memory. *Journal of Verbal Learning and Verbal Behavior,* 1971, **10,** 400–408.

Smith, A. D., D'Agostino, P. R., & Reid, L. S. Output interference in long-term memory. *Canadian Journal of Psychology,* 1970, **24,** 85–89.

Smith, E. E. Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6): *Linguistic functions in cognitive theory.* Hillsdale, NJ: Erlbaum, 1978. Pp. 1-56.

Tulving, E. The effects of presentation and recall in free recall learning. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6,** 175–184.

Tulving, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972. Pp. 382–404.

Tulving, E. Cue-dependent forgetting. *American Scientist*, 1974, **62,** 74–82.

Tulving, E., & Pearlstone, Z. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 381–391.

Watkins, M. J. Inhibition in recall with extralist 'cues'. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14,** 294-303.

Waugh, N. C. Presentation time and free recall. *Journal of Experimental Psychology*, 1967, **73,** 39–44.

Williams, M. D. *Some observations on the process of retrieval from very long-term memory*. Doctoral dissertation, University of California, San Diego, 1977.

MEMORY-BASED REHEARSAL

Ronald E. Johnson

PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA

I. Introduction

Within the last 10 years rehearsal increasingly has been assigned a major role in explanations of learning and retention. The ubiquity of the explanatory concept is evident by its usage in practically all experimental paradigms. To cite but a single study in each paradigm, one finds rehearsal interpretations attempted for free recall (Ashcraft, Kellas, & Needham, 1975), cued recall (MacLeod, 1975), backward recall (Buschke & Hinrichs, 1968), discrimination learning (Levin & Ghatala, 1976), serial recall (Wickelgren, 1967), serial probed recall (Ferguson & Bray, 1976), recognition (Jacoby, 1973), continuous recognition (Santa & Ranken, 1972), running memory span (Frank & Rabinovitch, 1974),

THE PSYCHOLOGY OF LEARNING
AND MOTIVATION, VOL. 14

ISBN 0-12-543314-X

Sternberg task (Seamon & Wright, 1976), paired associate learning (Izawa, 1976), concept identification (McVaugh, 1973), problem solving (Olton & Johnson, 1976), release from proactive inhibition (Schendel, 1976), Brown–Peterson task (Dark & Loftus, 1976), directed forgetting (Bjork & Woodward, 1973), imitative and observational learning (Bandura & Jeffery, 1973), motor learning (Stelmach & Bassin, 1971), incidental learning (Postman & Kruesi, 1977), and animal learning (Wagner, Rudy, & Whitlow, 1973). Rehearsal also has played pivotal roles in explanations of specific learning outcomes such as one-trial learning (Lockhead, 1961), reduced recall of once-presented units (Fritzen, 1975), trade-offs in remembering (Bellezza & Walker, 1974), spacing effect (Elmes, Sanders, & Dovel, 1973), feedback influence (Sassenrath, 1975), serial position effects (Rundus, Loftus, & Atkinson, 1970), von Restorff effect (Detterman, 1975), reminiscence (Woodworth & Schlosberg, 1954), amnesic effect (Detterman & Ellis, 1972), temporal and frequency discriminations (Shaughnessy & Underwood, 1973), and organization and clustering in recall (Weist, 1972).

The heightened interest in rehearsal perhaps stems from increased awareness that a complete account of the learning process requires knowledge of learners' cognitions during learning and recall. Concomitantly, the invocation of rehearsal explanations probably has been provoked by the intuitive validity of such notions, the tempting ambiguity of the explanatory construct, and the lack of knowledge regarding the conditions under which rehearsal influences performance. As a consequence, rehearsal explanations often have been post hoc incantations used to account for gains or losses in performance that otherwise were difficult to explain. The occurrence of an unfilled temporal interval similarly has been taken as prima facie evidence that rehearsal has occurred (e.g., Brown, 1958; Cermak, 1972; Hinrichs & Grunke, 1975; Jahnke & Davidson, 1967; Lockhead, 1961; Postman & Warren, 1972; Poulton, 1977).

Given the increasing ascendency of rehearsal, the time appears appropriate to examine its current conceptual status. A second purpose is to review critically selected topics in the rehearsal literature. Various conceptual and methodological issues are examined first. Then, in an examination of the comparability of outcomes across rehearsal experiments, assessments are made of the influence of three temporal variables on the effectiveness of rehearsal. Attention then focuses on the possible functions served by rehearsal. Evidence also is sought on processes that might differentiate rehearsal from other activities. Next, the use of rehearsal as an explanatory variable is examined through a detailed look at the literature on directed forgetting. Finally, an appraisal is made of various theoretical positions regarding the effects of rehearsal.

DEFINITION

The scientific usage of "rehearsal" has been heavily influenced by its usage in everyday language. Thus the *American Heritage Dictionary* (Morris, 1975) defines "rehearsal" as:

> 1. The act or process of practicing in preparation for a performance, especially for a public performance. 2. A verbal repetition or recital.

English and English's (1958) *Dictionary of Psychological and Psychoanalytical Terms* lists a similar entry:

> 1. Performing an act prior to the time when it will be needed; hence, 2. going over in one's mind previously studied data; repetition with a view to later recall. It may be involuntary, self-imposed, or other-imposed.

Both definitions, as well as much of the current literature on rehearsal, emphasize the factor of repetition. Repetition is a frequently used strategy in rehearsal, but the two concepts are not synonymous. The rehearser may engage in attempts at recoding previous learnings, or in attempting to develop additional retrieval cues, but the activity need not include repetition. A definition equating rehearsal with repetition thus is overly restrictive.

Rehearsal also needs to be differentiated from external reexposures to the task. Such reexposures are simply additional learning trials, and there is no merit in dubbing such presentation trials as rehearsal. Even when the reexposure occurs after a delay interval, as in *review,* the reexposure does not constitute a rehearsal period. A review typically does include selective reexposures to the original content, and/or presentations of new facts and generalizations. As defined here, however, such external presentations would not constitute rehearsal even though such learning activities do prepare learners for future performances. All learning experiences may be viewed as preparation for future performances, but if rehearsal is to be a useful concept, the term should designate some particular subset of preparatory learning activities.

Perhaps the most useful delimitation is to restrict usage to those preparatory learning activities that occur when the reexposures and/or new encodings are internally generated by the learner rather than by additional presentations external to the learner. Thus, rehearsal refers to preparatory learning activities performed upon the contents of the rehearser's own memory. Except for the limitation on the source of the content to be rehearsed, the "preparatory learning activities" that constitute rehearsal are otherwise left unrestricted. Covert mental practice is one particularly interesting form of rehearsal, but rehearsal also may consist of overt verbalizations as well as overt physical activity.

The preceding definition of rehearsal encompasses most of the previous investigations labeled as studies of rehearsal. A notable exception is the exclusion of studies in which learners overtly or covertly repeat a word during the exposure of that word. Such activity, however, is classified as rehearsal if the repetitive activity is based upon the learner's remembering of the word rather than upon additional external exposures to the learner. As the review proceeds, it will become clear that many of the reviewed studies have included rehearsal conditions, nonrehearsal conditions, and conditions combining various blends of rehearsal-nonrehearsal learning activities. Some investigators thus have included conditions in which some rehearsers are given additional access to a portion of the experimental materials (e.g., Johnson, 1970). Through such manipulations, it is possible to determine the extent to which rehearsers' success is limited by failures in remembering.

II. Methodological Issues

In this section of the article, attention is directed to several methodological problems that have long plagued researchers. The initial segment focuses on the assessment of covert rehearsal, while the second part is an examination of techniques used in preventing rehearsal. As will become evident, the proposed solutions for these two problems are not entirely satisfactory.

A. ASSESSMENT OF COVERT REHEARSAL

Researchers typically have assumed that learners rehearse whenever the opportunity occurs, either during presentation or retention (e.g., Buschke & Kintsch, 1970; Shaffer & Shiffrin, 1972). Whether rehearsal is actually occurring, however, is uncertain, and a number of researchers have attempted to assess covert rehearsal through various indices. Ulich (1967) recorded changes in electromyographic potentials (EMGs), pulse rate, and respiratory frequencies during the mental practice of a motor performance. Mean variations in the physiological measures differentiated the mental rehearsal group from groups engaged in active practice or observational learning, but some rehearsers did not show physiological changes during rehearsal. Further, rehearsers showing moderate levels of change in their EMG tracings showed better performance. Such outcomes complicate the task of experimenters desiring to use EMG recordings as an index of rehearsal, and further complication is evident in Locke and Ginsberg's (1975) investigation of EMGs arising from the rehearsal of alphabetical letters. Although muscular potentials were useful in rehearsal detection, speech EMG tracings were difficult to distinguish from those of nonspeech mouth movements. Lipreading judgments were found to be better indicants of rehearsal, but neither lipreading nor electromyography is likely to prove useful in detecting the rehearsal of material in which mouth movements are less distinct.

Pupillary size typically has been interpreted as an index of mental effort, and hence might index covert rehearsal in certain experimental contexts (Kahneman & Wright, 1971). So far, however, the autonomic measure has been used only in conjunction with covert rehearsal, and the argument would be stronger if it could be demonstrated that pupillary changes also covary with overt rehearsals. Pupillary responses are known to be influenced by many other variables (Kahneman & Wright, 1971), and alternative explanations of the covariation with rehearsal are possible.

Other researchers have requested learners to press a button whenever a rehearsal occurs (Ellis & Montague, 1973; Kroll & Kellicutt, 1972; Rowe 1975; Schendel, 1976). As Montague, Hillix, Kiess, and Harris (1970) note, however, such a requirement may induce extra rehearsals. The technique also assumes that the learner possesses awareness of rehearsals, is willing to report rehearsals that may have been forbidden, and will not lapse into periods of nonreporting.

Pausal times during self-paced learning also have been used as an indicant of rehearsal (Belmont & Butterfield, 1971). Introspective reports, gathered at the end of the testing session, indicated considerable correspondence between the subjective groupings of units and the specific locations of pauses. Response latencies during recall also showed some relationship to the pausal patternings during learning (Butterfield & Belmont, 1971). The durations of the pauses were not closely related to the percentages of correct recall, but this finding may be a consequence of learners adopting a strategy of not rehearsing those terminal units that are readily available from short-term memory. As the learner gains additional experience with a particular learning task, the evidence also indicates self-allocated changes in the length of maximum rehearsal pauses, change of the serial position at which maximum rehearsal occurs, and increases in the number of nonrehearsed units in the terminal positions (Belmont & Butterfield, 1971).

Some investigators have assumed that rehearsal activity during the retention interval may be indexed by the accuracy of performance on a rehearsal-preventive task (Kroll & Kellicutt, 1972; Reitman, 1971, 1974; Roediger & Crowder, 1972; Tarpy, Glucksberg, & Lytle, 1969). Roediger and Crowder, for example, compared three groups that were differentially motivated to rehearse (''remember and rehearse,'' ''remember without rehearsing,'' and ''forget''), and found that performance on the interpolated task was inversely related to the incentive to rehearse. Variations in group performances on such interpolated tasks, however, typically are small in comparison with the differing amounts of time that are presumed to be spent in rehearsing. A further complication is that individual differences in abilities to perform such prevention tasks may lead to misleading inferences regarding the occurrence or nonoccurrence of rehearsal (also see Section II,B).

Rather than suffer the methodological ambiguities of covert rehearsal, researchers increasingly have turned to overt rehearsal as an alternative methodology. Given the potential advantages of directly observing rehearsal, an important question is the equivalence of overt and covert rehearsal. A search of the literature located seven studies of memory-based rehearsal in which comparison was made of groups that differed only in whether rehearsal was overt or covert. Three of these studies reported greater postrehearsal remembering for covert responding (Jeffery, 1976; Murray, 1967; Reynolds, 1967), two showed nonsignificant differences (Johnson, 1968; Roenker, 1974), and two reported superiority for overt rehearsal (Izawa, 1976; Whitten & Bjork, 1977).

Detailed comparisons of the methodologies of these seven studies, unfortunately, offer little insight into the conditions under which the two types of rehearsal lead to different outcomes. There is no consistent ordering of outcomes as a function of whether the paradigm was paired-associate learning (Izawa, 1976; Johnson, 1968; Reynolds, 1967) or free recall (Murray, 1967; Roenker, 1974). The extent to which rehearsers were free to adopt their own rehearsal strategies (Jeffery, 1976; Roenker, 1974), as opposed to a formal structuring of the rehearsal activity by the experimenter (Izawa, 1976; Johnson, 1968; Murray, 1967; Reynolds, 1967; Whitten & Bjork, 1977), also appears unrelated to overt–covert differences. The differential outcomes are not obviously related to the level of task proficiency at the time that rehearsal began, nor to the extent to which rehearsal was based on short-term rather than long-term memory. Until there is an unraveling of the conditions under which the two types of rehearsal lead to equivalent outcomes, it cannot be assumed that overt rehearsal is a mirror image of covert rehearsal.

B. PREVENTION OF REHEARSAL

The control of rehearsal is widely recognized as a serious methodological problem, and experimenters have used various techniques to minimize or eliminate illicit rehearsal. Such control techniques typically lead to lower recall, but whether such losses are a direct consequence of eliminating rehearsal is always open to question. Changes in the level of recall, unfortunately, may provide a misleading index of the occurrence of rehearsal. Although the practice is exceedingly common, the presence of beneficial rehearsal cannot be assumed simply because recall was relatively good. Similarly, poor recall does not necessarily signal the absence of rehearsal (Götz & Jacoby, 1974; Jacoby & Bartz, 1972; but see Dark & Loftus, 1976).

1. Rapid Presentation Rates and Time-Sharing Tasks

To prevent rehearsal during learning, some investigators have used fast presentation rates that presumably allow little time for rehearsal (Cohen, 1967; Lockhead, 1961; Pollack, 1963). Fast presentation rates, however, do not preclude the possibility that the learner might rehearse a portion of the content and not the rest (Cohen, 1967). This concern is less applicable to methodologies in which the learner's rehearsal time is minimized by a requirement to perform some concurrent task during presentation (e.g., Cohen, 1973). When learners are required to switch tasks, however, the simple entry of new content into processing can introduce interference (Sternberg, 1969).

2. Interpolated Rehearsal-Preventing Tasks

Attempts have been made to find rehearsal-preventing tasks that engage learners' processing capacities but yet do not act as direct sources of interference. Learners typically are required to monitor or transform stimuli in such interpolated tasks, but are not required to remember the content. Despite their apparent neutrality, however, such tasks may interfere with remembering. Neimark, Greenhouse, Law, and Weinheimer (1965), for example, found that low meaningful CVCs were remembered less well when the interpolated task was one of reading low meaningful CVCs rather than medium or high CVCs. Recall of high meaningful lists, in contrast, was not differentially affected by the meaningfulness of the interpolated task.

In an experiment by Brown (1958), subjects were presented pairs of consonants to be learned, and then given either an unfilled temporal interval or else required to read pairs of digits. The imposition of the reading task severely restricted the amount of recall. When the interpolated reading was content similar to that which was to be remembered, recall was slightly less than when the interpolated reading was dissimilar content (25.6 vs 30.6% $p<.05$). These results seemingly argue that it was the prevention of rehearsal rather than interference from the interpolated task that was primarily responsible for the decline. Yet, when the temporal interval between learning and interpolated reading was either .78, 2.34, or 4.68 sec, the corresponding proportions recalled were .41, .54, and .59. These figures compare with 94% recall in the earlier experiment in which recall was immediate, and there was no interpolated reading. Thus, increasing rehearsal time from .78 to 4.68 sec prior to the interpolated reading only moderately reduced the interference generated from the seemingly neutral task of reading.

Using methodology similar to Brown's (1958), Pollack (1963) reported that recall was worse when learners were required to read more interpolated digits during the rehearsal-prevention task. Slower rates of presentation of the interpolated digits also resulted in less interference than faster rates. When learners were permitted to delay their oral reporting of the interpolated digits, recall of the desired sequence was better than when the oral reporting was required immediately. Across the various experimental conditions in Pollack's experiments, the mean levels of recall ranged from approximately 30 to 95%. An interpolated task thus can create considerable interference with recall even when learners are not required to learn the interpolated materials.

Given the wide variety of tasks used to prevent rehearsal, there is clear need for information regarding task dimensions that govern rehearsal control. Existing evidence indicates that the difficulty of the interpolated activity is inversely related to recall of the original learning. Variations in the difficulty level of the preventive activity are widely assumed to influence recall by differentially restricting the opportunity to rehearse, but an equally plausible explanation is that such interpolated tasks generate differing levels of interference. Either explanation thus could account for the fact that interpolated tasks requiring arithmetical calculations result in less retention of the original learning than does the reading of numbers (Dillon & Reed, 1969; Murray, 1966; Peterson, 1969). Similarly, the difficulty levels of interpolated card sorting tasks were inversely related to subsequent recall (Murdock, 1965). In the remembering of a list of words, Petrusic and Jamieson (1978) found recall was best when an interpolated interval was left unfilled, and increasing amounts of forgetting were found with interpolated tasks of listening to instrumental music, listening to vocal music, and shadowing.

Difficulty levels of interpolated tasks have been determined primarily by intuition, but Peterson (1969) has provided empirical data suggesting a three-level hierarchical ordering of concurrent (or interpolated) tasks according to the degree of attention required. From lowest to highest attentional requirements, the levels are: (*a*) *emission* of a previously learned performance, e.g., reciting the alphabet, (*b*) *reproduction* of a sequence of external stimuli, e.g., reading aloud, and (*c*) *problem-solving* transformations of input stimuli, e.g., solving anagrams. Peterson's ordering is a useful start toward the classification of rehearsal-prevention activities, and the literature on "levels of processing" affords additional insight into task dimensions that may influence the frequency of unauthorized rehearsal (e.g., Cermak & Craik, 1978).

3. Recurrent Subtraction

Perhaps the most widely used method of preventing rehearsal is requiring learners to "count backward" by recurrently subtracting a designated number. Most experimenters have followed Peterson and Peterson's (1959) lead in having learners subtract by 3s, although other investigators have required subtractions of 7s (Flexser, 1978; Kroll & Kellicutt, 1972), 2s (Cuvo, 1974), and 1s (Santa & Ranken, 1972). When iterative subtraction is required, learners typically show poorer remembering than when the delay interval is left unfilled (e.g., Peterson & Peterson, 1959), and recall appears to be inversely related to the number of subtractions performed during the retention interval (Roediger & Crowder, 1972). Even backward counting by 1s resulted in a substantial decline in memory performance (Santa & Ranken, 1972). Roediger and Crowder (1972) argued that self-paced subtraction is a more effective rehearsal preventative than experimenter-paced subtraction, but the issue is one that requires an empirical answer.

Kroll and Kellicutt (1972) reported poorer remembering of trigrams after subtracting 7s than 3s. When learners were required to indicate rehearsals by pressing a button, more rehearsals were reported during the subtraction by 3s. Even though learners were instructed that their order of priority should be backward counting, the reporting of rehearsals, and then the recalling of memory letters, the learners averaged approximately one rehearsal during each 9-sec retention interval. Thus, there appears to be some surreptitious rehearsal during subtraction performances although the extent to which such rehearsals were induced by the reporting requirement is uncertain.

4. *Signal Detection*

Signal detection has been used as a task that might suffice in preventing rehearsal, but yet would not serve as a source of interference in remembering. As would be expected, decrements in recall following signal detection are related to the similarity of content of the detection task to that of the original learning (Reitman, 1971, 1974). More interesting, however, is that when there is no similarity of content between the learning task and the detection task, the length of time spent in signal detection is unrelated to the subsequent level of recall (Atkinson & Shiffrin, 1971; Reitman, 1971, 1974; Shiffrin, 1973). In contrast, when signal detection was followed by an arithmetic task, there was a considerable decline in recall (Atkinson & Shiffrin, 1971; Shiffrin, 1973). Does this mean the arithmetic interfered with remembering, or does it mean that surreptitious rehearsal was occurring during signal detection? Since an instruction to rehearse after each signal detection leads to higher performance than signal detection alone (Atkinson & Shiffrin, 1971), the results suggest either than no rehearsal or else nonoptimal rehearsal occurred during signal detection.

In a well-reasoned series of experiments, Reitman (1971) and Shiffrin (1973) concluded that their learners had not engaged in rehearsal during signal detection. Since this series incorporated detailed analyses to detect rehearsal, it is instructive to examine their evidences. Aside from noting postexperimental self-reports that learners had no difficulty in avoiding rehearsal, Reitman (1971) also reported that detection accuracy and latency did not differ on trials requiring memory of the words as opposed to control trials on which learners knew that memory would not be required. Further, since rehearsal by individual learners conceivably could have been masked by group means, Reitman also compared experimental and control detection scores for each learner, but a statistical test for ''outliers'' also revealed no significant discrepancies.

Shiffrin's (1973) evidence against rehearsal during signal detection included the results of learners who overtly rehearsed a single time after each detection. Since memory improved, but did not do so in an earlier experiment, Shiffrin argued that learners had not been rehearsing in the earlier experiment. In addition to reporting antirehearsal evidences comparable to Reitman's (1971), Shiffrin noted that learners felt it would have been silly to rehearse since they always remembered the material without rehearsal. The likelihood of rehearsal probably was further reduced by the provision of monetary rewards for accuracy in signal detection but not for accuracy in recall.

Reexamining the issue, however, Reitman (1974) concluded that surreptitious rehearsal does occur during signal detection. The 1974 study included three within-subjects conditions: (*a*) an antirehearsal condition in which learners were requested to avoid rehearsing during signal detection, (*b*) a rehearsal condition in which subjects were told to rehearse as often as they could without allowing a decline in detection performance, and (*c*) a control condition of signal detection without any memory requirements. Since signal detection accuracies and latencies were worse in the rehearsal than the control condition, Reitman concluded that rehearsal does disrupt detection performance. An alternate interpretation, however, might be that the rehearsers unwisely allocated too much attention to rehearsal and too little to signal detection. A comparable analysis, in fact, showed no difference between the antirehearsal and the control conditions of the 1971 study. According to Reitman, however, the outlier test might have been insensitive if too many actual rehearsers had contributed difference scores that inflated the variance pool. As a more sensitive method of detecting rehearsal, Reitman (1974) proposed seven within-subjects comparisons of signal detection and/or recall. Two measures were based on the assumption that better detection performance in the no-memory control condition, as opposed to an experimental condition, was evidence of rehearsal. Three other measures were based on the assumption that trade-offs in signal detection and rehearsal would be evident in negative correlations between detection performance and recall. Performing well in recall but poorly in detection thus was assumed to be evidence of rehearsal. The final two measures were based on the assumption that greater variance in detection performance across five successive 3-sec blocks of detection in the experimental condition, as compared with the control condition, reflected shifts of attention from detection to rehearsal. Each subject was judged to be rehearsing if any of the seven measures indicated worse signal detection performance on the memory-load trials than on the control trials. When judged against these seven measures, 25 of the 29 learners showed at least one piece of evidence that their detection performance was worse under the rehearsal condition than the control condition. Under instructions to avoid rehearsal, 23 of the 29 learners nevertheless were judged to have been rehearsing.

Interpreting Reitman's (1974) results, however, presents some difficulty. First, it is not obvious that poor signal detection is necessarily a sign of rehearsal. Across all learners, a decrement in detection under the memory-load conditions could be due to the forced time sharing of processing capacity required for both storage and detection rather than to active rehearsal. Also questionable is the technique of examining each learner's performances against multiple screening criteria and then assuming that one suspicious index is proof of rehearsal. As the number of indices increase, there is increased probability that chance alone would erroneously allow classification of a nonrehearser as a rehearser. The probability of such misclassifications will depend upon the reliability of the classification measures, and some of the measures undoubtedly allow unreliable estimates. Indeed, though 23 of the 29 subjects in Reitman's avoid-rehearsal condition were classified as rehearsers, the mean performances of the 29 subjects as a group on the seven rehearsal indicators do not indicate levels of rehearsal greater than that of chance. Reitman's argument is weakened further by the lack of a statistical test comparing signal-detection performances in the rehearsal and antirehearsal conditions. Assuming some cooperation from learners, there should be less rehearsal in the antirehearsal condition, and measures for classifying learners as rehearsers or nonrehearsers should reflect this difference. Finally, the validity of the seven measures in accurately classifying rehearsers is not evident by the levels of recall shown in Reitman's (1974) Table 2. There appear to be no differences in recall as a function of whether Reitman's measures classified the learners as rehearsers or nonrehearsers. Whether learners were instructed to rehearse or to avoid rehearsal, however, was an important determinant of recall performance. Those who were instructed to avoid rehearsal were significantly poorer in recall. Contrary to Reitman's assertion, then, the application of the seven criteria to the antirehearsal subjects does not necessarily provide "strong behavioral evidence of rehearsal."

5. *Incidental Learning Instructions*

Another approach to controlling rehearsal has been to mislead learners into thinking there will be no later test of retention. Underwood and Postman (1960) attempted to conceal the purpose of a return to the laboratory by telling learners they would be retested on an additional task. Despite the attempt to mask the purpose of the retest, 16 out of 144 learners admitted to rehearsing at least three times orally or else writing the verbal units at least once during the week-long retention interval.

To eliminate rehearsal during the learning of a list of words, Marshall and Werder (1972) also used an incidental learning format in which learners were required to assign specified numbers to each vowel and then compute the sum. Other learners were required to perform the coding task but also were requested to learn the words. Only one learner out of 80 admitted to rehearsal, and there were no overt indicators of rehearsal such as lip movements.

6. *Instructions to Avoid Rehearsal*

Experimenters frequently request learners to avoid rehearsal (e.g., Roediger & Crowder, 1972; Shaughnessy & Underwood, 1973; Weiner & Reed, 1969), but there have been surprisingly few experimental comparisons of the effectiveness of such admonitions. Underwood and Keppel (1962) instructed some learners to avoid rehearsal during a 24-hr retention interval. In postexperimental questioning of 64 learners, two admitted to having engaged in systematic rehearsal. For learners not given antirehearsal instructions, 16 of 64 reported systematic rehearsal. Requesting learners to avoid rehearsal thus appears to lower the incidence of rehearsal. An alternate possibility, however, is that the learners given antirehearsal instructions are under greater social pressures to conceal illicit rehearsal. Statistical analyses of differences in later recall, in fact, showed no main effects for the antirehearsal manipulation.

Brodie and Prytulak (1975) instructed various groups of learners to allocate rehearsals to words either in the beginning, middle, or end of their list and to avoid rehearsing the other words. The portion of the list that subjects were instructed to rehearse differed according to which portion was being presented, and the learner was required to furnish a continual oral report of any word he was thinking about. Despite the complex instructions, Brodie and Prytulak found "almost perfect compliance with instructions to limit overt rehearsals to specific items," but postexperimental questioning revealed that 36 of 48 subjects had violated instructions by covertly rehearsing forbidden words. In a subsequent study learners were instructed to think about and overtly rehearse only the word being presented, even if such a strategy lowered overall learning. Following recall, learners were given a list of the words and requested to indicate any other words that they thought about during the time a particular word was being presented. All subjects reported covert violations of the instruction to limit rehearsal to the item being presented. The mean number of violations per list was 14.8, $\sigma = 11$, $R = 3$ to 44. Since subjects accurately reported that their *overt* rehearsals had been restricted to the word that was being exposed, Brodie and Prytulak assumed that the reports of covert rehearsal also were correct. However, since the covert reports were taken after recall, the possibility exists that learners' reports were influenced by the differential recall strengths or familiarity of the items. The reporting situation also may have induced demand-characteristics to find associative linkages that had been "rehearsed." Taking the results at face value, however, the outcome suggests that considerable surreptitious covert rehearsal occurs even when the learner is required to rehearse overtly.

In summary, learners frequently initiate rehearsal during unfilled retention intervals. The rehearsal activity, however, does not always lead to increments in recall. Rehearsal prevention techniques typically lead to lower levels of recall, and there is some knowledge regarding the task dimensions associated with lower recall. A recurrent problem, however, is the possibility that the interpolated task acts as interference rather than as a rehearsal preventive. Assessing the effectiveness of such rehearsal control techniques requires a valid index of covert rehearsal, but each of the existing indices has possible flaws. Since there is no infallible index, the researcher probably is well advised to obtain several concurrent measures of rehearsal activity. An alternate route to controlling rehearsal activity might be to present the learning task in an incidental learning paradigm or to select learning tasks that are less susceptible to effective rehearsal.

III. Empirical Issues—Temporal Influences

Provided that rehearsers have reason to expect a delayed test, there is a strong relationship between the relative frequencies of overt rehearsal and later postrehearsal performance (Rundus, 1971; Rundus & Atkinson, 1970). Given the generality of this finding across different experimental contexts (e.g., Brodie & Prytulak, 1975; Johnson, 1975; Kroll, Kellicut, & Parks, 1975; Murdock & Metcalfe, 1978), the question occurs as to whether other empirical consistencies exist across rehearsal experiments. To assess the likelihood of such regularities, the literature was searched to ascertain the influence of three temporal parameters on the effectiveness of rehearsal.

A. DURATION OF REHEARSAL

Many investigators have varied the length of the rehearsal period either by changing the interitem interval or else the time between original presentation and the criterion task. As might be expected, longer rehearsal periods typically result in higher levels of performance (Hockey, 1973; Leicht, 1968; Longstreth, 1971; McDaniel & Masson, 1977; Meunier, Kestner, Meunier, & Ritz, 1974; Meunier, Ritz, & Meunier, 1972; Monty, Karsh, & Taub, 1967; Moss & Sharac, 1970; Nodine, Nodine, & Thomas, 1967; Penney, 1975; Peterson & Peterson, 1959; Peterson, Peterson, & Ward-Hull, 1977; Peterson, Thomas, & Johnson, 1977; Pollatsek & Bettencourt, 1976; Reed, 1970; Roediger & Crowder, 1972;

Roenker, 1974; Sackett, 1935; Sanders, 1961; Smyth, 1975; Starr, Harris, & Aronoff, 1970; Ternes & Yuille, 1972).

As the rehearsal period is lengthened, the temporal interval also lengthens between the original learning and the final criterion performance. Despite the confounding of rehearsal duration with retention interval, there have been no reports of inferior performances after longer rehearsal periods. Approximately a third of the studies, however, have shown no differences in performance as a function of rehearsal duration (Cermak & Levine, 1971; Colegate & Eriksen, 1970; Ellis, 1970; Glenberg, Smith, & Green, 1977; Kestner & Walter, 1977; Meunier, Meunier, & Ritz, 1971; Meunier *et al.*, 1972; Ryan, 1969; Shaffer & Shiffrin, 1972). In some instances, there are plausible reasons for the failure of duration to have an influence. For example, neither Cermak and Levine (1971) nor Meunier *et al.* (1971) instructed subjects to rehearse during the interval, and it is uncertain whether the learners availed themselves of the opportunity. Other failures may have been due to the brevity of the longer rehearsal duration (e.g., 2.5 sec in Colegate & Eriksen, 1970), small *N*s and large variability (Ellis, 1970), ceiling limitations (Meunier *et al.*, 1972), and low levels of learning across conditions (Kestner & Walter, 1977). For the remaining experiments, however, there is no obvious explanation, and it appears that increased amounts of rehearsal do not necessarily lead to increased recall.

Even when increased rehearsal duration leads to superiorities in immediate recall, the effects may not persist at longer retention intervals either in recall (Meunier *et al.*, 1972; Nodine, 1969, $p<.08$; Reed, 1970) or in recognition (Roediger & Crowder, 1972). In Roenker's (1974) study, the groups having an initial recall advantage from additional rehearsal were actually inferior on a final free recall.

The potential durability of gains acquired in extended rehearsal periods is evident in the relatively greater advantages of longer rehearsal in surmounting periods of interpolated interfering activities (Pollatsek & Bettencourt, 1976; Sanders, 1961). Similarly, McDaniel and Masson (1977) found that the initial recall advantages from longer interitem rehearsal intervals persisted across a 24-hr retention interval.

Several investigators have reported that the earlier portions of the rehearsal interval are more critical than the later phases (Estes, 1973; Meunier *et al.*, 1974; Nodine, 1969; Sackett, 1935). When two separate rehearsal periods are allocated, however, the second period appears to be no less important than the first one. Following each of two presentations, learners in Pollatsek and Bettencourt's (1976) experiment had various lengths of rehearsal time and interference-activity time. Longer rehearsal times resulted in superior recall, but the beneficial effects of additional rehearsal time were equivalent regardless of whether added to the first or to the second rehearsal period. In a similar manipulation, however, DeRemer and Agostino (1974) found performance was better when the longer rehearsal period was allocated to the second presentation rather than the first. This outcome supported their view that a spacing effect occurs only when there is sufficient rehearsal time to allow a different encoding on the second presentation. The explanation for the differing results of the two experiments is not obvious since the studies differed in the content to be remembered, the presence or absence of an interference task, and other methodological procedures.

Whether an extended rehearsal period influences remembering depends upon the nature of the processing during the extended period. Peterson and Peterson (1959) found no improvement in performance when learners were given extra time in which covert rehearsal could have occurred, but the extra time was beneficial when the learners engaged in overt repetitive rehearsal. The importance of the rehearsal activity in mediating the influence of increased rehearsal is further illustrated in Hockey's (1973) work. When learners were instructed to use an active strategy of grouping and rehearsing, the longer rehearsal durations that were available during slower rates of presentation led to superior performances. When learners engaged in a passive reception strategy, longer rehearsal durations resulted in poorer performances.

The influence of rehearsal duration also appears related to the type of content that is rehearsed. Ternes and Yuille (1972) allowed different durations of rehearsal prior to the attempted recall of visually presented line drawings of familiar objects or else visually presented names of the same objects. The remembering of the words, but not the pictures, was enhanced by longer rehearsal periods. Increases in rehearsal duration similarly had no influence on pictorial recognition (Shaffer & Shiffrin, 1972). For these last two studies the presentation modality was visual, and performances were based upon long-term memory. After reviewing the literature on the effects of stimulus modality on short-term memory, Penney (1975) concluded that visual stimuli are recalled better when the presentation rates are sufficiently lengthy to allow rehearsal and recoding. With auditory stimuli, however, the presentation rate either had no effect, or else the faster rate *improved* recall provided that the recall requirement was for serial recall rather than free recall. Penney's conclusions thus suggest that presentation modality and the recall requirement both influence the extent to which rehearsal duration influences performance on short-term memory tasks.

B. DELAY BETWEEN LEARNING AND REHEARSAL

The length of the temporal delay between learning and rehearsal seemingly should influence the effectiveness of rehearsal. With longer retention intervals prior to rehearsing, more forgetting would occur, and the rehearser presumably would have less to rehearse. Spaced presentations, however, often lead to better performance than massed presentations (e.g., Melton, 1970; Underwood, 1970). To the extent that a rehearsal is equivalent to a second presentation, a delayed rehearsal might be expected to be more effective than an immediate rehearsal (Whitten & Bjork, 1977). Additional advantage could accrue from strengthening the existing rememberings shortly before the criterion task.

Only two studies were found in which there was systematic manipulation of the delay interval prior to rehearsal, and the results of both suggest that rehearsal is less effective if delayed too long. Bandura and Jeffery (1973) examined the influence of coding and rehearsal on the observational learning of modeled movements. Prior to observing complex movement sequences, some learners were taught either a patterned or an arbitrary memory code that could be used in labeling component responses. After each modeled movement, learners engaged either in symbolic overt rehearsal of the memory code, motor overt rehearsal of the modeled patterns, or a rehearsal-impeding task of signal detection. Following the entire series of demonstrations, learners engaged either in delayed rehearsal or in signal detection. Learners then attempted to reproduce the modeled performances. The results showed that coding and immediate rehearsal both were essential for superior remembering. Immediate symbolic rehearsal facilitated remembering only when there had been coding during acquisition. The effects of motor rehearsal, however, were small and limited to those learners receiving no formal code and those receiving the arbitrary symbolic codes. Although immediate rehearsal markedly facilitated later performance, delayed rehearsal had no influence on imitative reproduction.

Whitten and Bjork (1977) varied the interval between the initial presentation of a noun pair and a later rehearsal or an overt attempt at recall. Following a 2-sec initial presentation of the noun pair, learners engaged in 4, 8, or 14 sec of shadowing prior to a 3-sec rehearsal period or test trial, and then engaged in an additional 13, 9, or 3 sec of shadowing to complete the 22-sec series of events. After a block of 12 such trials, learners attempted free recall of all words in the block. Longer delays in the interval between initial presentation and the occurrence of the test trials resulted in improved end-of-block retention. For the rehearsal trials, however, there was a nonsignificant 6% decline in recall as the rehearsal period was delayed from 4 to 14 sec. The rehearsal conditions thus did not show a spacing effect, and the outcome suggests that delay prior to rehearsing may lessen the advantages of rehearsing.

C. PERSISTENCE OF GAINS FROM REHEARSAL

An important question is whether the learnings achieved through rehearsal persist beyond the initial test of performance. Accordingly, a number of experimenters have followed the initial test of performance with a delayed criterion test, or else have varied the length of the retention interval prior to the initial test of performance. The delayed criterion tests have been administered as soon as 10 sec later (Ferguson & Bray, 1976) or as long as 7 months later (Rubin-Rabson, 1941). In the majority of studies, the advantages of rehearsers on an immediate test of learning also were evident on a delayed retention test (Bandura, Jeffery, & Bachicha, 1974; Bellezza & Cheney, 1973; Brodie & Prytulak, 1975; Carter & Van Matre, 1975; Cuvo, 1974; Flexser, 1978; Jeffery, 1976; Jongeward, Woodward, & Bjork, 1975; MacLeod, 1975; Meunier *et al.*, 1974; Peterson, 1969; Peterson & Peterson, 1959; Rubin-Rabson, 1941, 1-week but not 7-month delay; Rundus *et al.*, 1970; Sanders, 1961; Wetzel, 1975; Wetzel & Hunt, 1977). For a smaller number of studies, the rehearsers' initial performance advantages did not persist across the delay interval (Ferguson & Bray, 1976; Jacoby, 1973; Meunier *et al.*, 1972; Nodine, 1969; Roenker, 1974). Finally, some experimenters have reported that the initial advantages of certain rehearsal groups persisted across a delayed retention interval whereas other groups of rehearsers lost their short-term advantages (Bellezza *et al.*, 1975; Bellezza & Walker, 1974; Bower & Reitman, 1972; Götz & Jacoby, 1974; Palmer & Ornstein, 1971; Thompson & Clayton, 1974; Tzeng & Hung, 1973).

Bandura and Jeffery's (1973) study was one of the few in which the rehearsal groups did not maintain their relative superiorities on a delayed test, but the experimenters attributed the failure to the forgetting of the arbitrary symbolic codes that mediated the rehearsers' initial superiorities. There is no obvious explanation for Nodine's (1969) rehearsers failing to maintain their performance advantages, but the differences at the delayed interval narrowly missed statistical reliability. In the remaining four experiments in which the initial advantages of rehearsers did not endure, the various experimenters noted a high likelihood that the rehearsers had engaged in maintenance-type rehearsal rather than pursuing semantic encodings (Ferguson & Bray, 1976; Jacoby, 1973; Meunier *et al.*, 1972; Roenker, 1974).

Learners' awareness of forthcoming rehearsal opportunities conceivably could influence the quality and durability of encodings. Götz and Jacoby (1974) found that learners who were precued that a rehearsal-prevention task would be required performed better on an unexpected final delayed test than learners who were precued that the retention intervals would be unfilled. As interpreted by Götz and Jacoby, the learners engaged in deeper semantic encoding when it was known the delay interval would be filled. Controlling for additional variables, however, Dark and Loftus (1976) found that subjects' awareness of whether rehearsal would be allowed or prevented had no influence on initial or final recall.

Awareness of a delayed test of memory also could induce attempts to form encodings that would bridge the temporal gap. In eight of nine experiments in which subjects were informed at learning of a forthcoming delayed test, or were made aware via a within-subjects manipulation of differing retention intervals, the superiorities of the rehearsal groups persisted across the retention interval (Bellezza & Cheney, 1973; Brodie & Prytulak, 1975; Carter & Van Matre, 1975; Meunier *et al.*, 1974; Peterson, 1969; Peterson & Peterson, 1959; Rubin-Rabson, 1941; Sanders, 1961; but not Ferguson & Bray, 1976). Even when unaware of a delayed test, however, rehearsers in 9 of 14 comparisons nonetheless maintained their superiorities at a delayed testing (Bandura & Jeffery, 1973, end of original session; Bandura *et al.*, 1974; Cuvo, 1974; Jeffery, 1976; Jongeward *et al.*, 1975; MacLeod, 1975; Rundus *et al.*, 1970; Wetzel, 1975; Wetzel & Hunt, 1977; but not the following: Bandura & Jeffery, 1973, 1-week delay; Jacoby, 1973; Meunier *et al.*, 1972; Nodine, 1969; Roenker, 1974). Bellezza and his colleagues also reported two studies in which superiority at a delayed retention interval was dependent upon rehearsers' awareness at learning that there would be a delayed test (Bellezza *et al.*, 1975; Bellezza & Walker, 1974).

In the five studies in which the benefits of rehearsal persisted for some rehearsal groups but not others, the groups in two studies were made aware that a delayed retention test would be administered (Bower & Reitman, 1972; Palmer & Ornstein, 1971), whereas the rehearsers in the other three experiments were not informed of a delayed test (Bandura *et al.*, 1974; Götz & Jacoby, 1974; Tzeng & Hung, 1973).

Overall, then, rehearsers who are aware of a forthcoming delayed test appear somewhat more likely to form encodings that persist beyond the immediate test of memory. Even so, the empirical literature provides ample evidence that awareness is not a prerequisite for enduring gains. The advantages evident in rehearsers' performances on their initial criterion test often persist even when the rehearsers are unaware of a delayed test of retention.

IV. Rehearsal Functions and Processes

Rehearsal has been portrayed as serving various functions in learning and retention. Although there have been few experimental attempts to separate the postulated functions, existing evidence suggests that rehearsal can serve various roles. The next section outlines several probable functions. In addition, evidence is sought on the question of whether there are processing facilities or capacities that are allocated exclusively to the rehearsal system. The question also is raised as to whether the process of rehearsal alters the quality of the memorial representations.

A. REHEARSAL FUNCTIONS

A widely held assumption is that rehearsal preserves or maintains input in short-term memory (Atkinson & Shiffrin, 1968; Broadbent, 1958), and the empirical literature provides ample evidence that maintenance can occur (e.g., Meunier *et al.*, 1972; Woodward, Bjork, & Jongeward, 1973). Less clear, however, is the effect of such maintenance on long-term retention. A common finding has been that maintenance rehearsal influences delayed recognition but not recall (Craik & Watkins, 1973; Glenberg *et al.*, 1977; Shiffrin, 1973; Woodward *et al.*, 1973). Other experiments, however, have shown that longer periods of maintenance rehearsal can result in gains in delayed recall (Baddeley, 1978; Dark & Loftus, 1976; Darley & Glass, 1975; Modigliani, 1976).

Evidence also suggests that rehearsal can increase the accessibility both of content already in memory (Sternberg, 1969; 1970, Exp. V) and content that is being learned (Lachman & Mistler, 1970). If rehearsal increases accessibility, the relationship between rehearsal and later recall should weaken when the parameters of the task minimize the importance of accessibility. Consistent with an accessibility interpretation, Einstein, Pellegrino, Mondani, and Battig (1974) found stronger relationships between overt rehearsal and free recall when the presentation of items was successive rather than simultaneous. Similarly, the relationship between rehearsal and recall was at its strongest during the early learning trials—a time when problems of accessibility presumably would be greatest.

Another probable consequence of rehearsal is that of increasing resistance to the effects of interfering tasks. As noted previously, Pollatsek and Bettencourt (1976) and Sanders (1961) found that longer periods of rehearsal were advantageous in surmounting interference arising from interpolated activities. Similarly, Pollack (1963) found that longer durations of rehearsal prior to an interfering task resulted in higher levels of recall. Additional comparisons showed improved recall when there were longer unfilled intervals of time between the interfering units. Such outcomes permit the interpretation that rehearsal can assist in allaying interference.

Rehearsal also provides learners opportunity to form new associations. As suggested by Neisser (1967, p. 240), ''rehearsal is not the invigoration of an old structure but the synthesis of a new one.'' Illustrating this possibility, Peterson *et al.*'s (1977) rehearsers correctly formed a new mental matrix of numbers by combining two matrices that had been memorized previously. Similarly, the superior postrehearsal performances of rehearsers in Johnson's (1975) study were shown to be due in part to the acquisition of new correct associations formed during the rehearsal period.

B. DEDICATED FACILITIES FOR REHEARSING

The rehearsal system apparently accesses processing capacities that are separate from those used by the sensory system (Turvey, 1966). Following the learning and rehearsal of an initial task, the sensory register was filled and tested via Sperling's (1960) partial report methodology. Learners then attempted recall of the rehearsed task. Performances on the Sperling task were equivalent regardless of the similarity of content between the two tasks and regardless of whether an initial learning task was even required. The circulation of rehearsed information thus had no influence on storage processes within the sensory system.

A related possibility is that the rehearsal system has separate storage or processing capacities that are exclusively dedicated to the rehearsal process. Traces formed during rehearsal thus might be stored in a different location from traces that are not rehearsed. If so, rehearsal should increase short-term memory capacity by freeing storage slots in the regular short-term store. In a test of this possibility, Sperling and Speelman (1970) found that rehearsal led to a slight *decrease* in estimated memory capacity. Thus, short-term memory traces formed during rehearsal apparently are not shunted to a separate storage area.

Working with retarded learners, McBane (1976) tested Fisher and Zeaman's (1973) contention that rehearsing information in one conceptual category (e.g., forms) would not interfere with the retention of information rehearsed from a second category (e.g., colors). Following the determination of capacity limits for each of the two types of content, learners received a longer list consisting of capacity loads of both form stimuli and color stimuli. Contrary to the hypothesis of independent capacities, the rehearsal of problems in one dimension interfered with remembering of the other dimension.

Posner and Rossman (1965), in contrast, assumed that the processing capacity available for rehearsal would vary as a function of other processing demands within a limited capacity system. When subjects were required to perform irrelevant cognitive transformations on the content to be remembered, such as adding the digits that were to be remembered, less processing capacity presumably was available for rehearsal. Supporting their interpretation, the greater the number or difficulty of required transformations, the poorer was retention. An additional experiment showed that the losses occurred primarily in content encountered immediately prior to the transformation. The digits involved in the transformation, in fact, were recalled somewhat better than would be expected from their serial positions.

Learners in Moss and Sanders' (1973) study received a signal during the presentation informing them to begin the additional task of classifying each new consonant as "early" or "late" in the alphabet. In apparent contrast with Posner and Rossman's (1965) finding that the transformed units were remembered well, the classification task in Moss and Sanders' experiment led not only to poorer remembering of the consonants encountered prior to the classification requirement but also to poorer remembering of the classified consonants themselves. As interpreted by Moss and Sanders, the classification task consumed time that normally would have been available for rehearsing the earlier consonants in the series.

Since rehearsal often is performed under severe time constraints, it is not surprising that a requirement to perform an additional task while rehearsing leads to poorer postrehearsal performances (e.g., Baddeley & Hitch, 1974). Less obvious, perhaps, is that rehearsal may lessen performances on a task that is performed concurrently. In several studies subjects learned and rehearsed a sequence of items in anticipation of a later test. During the acquisitional sequence, subjects also were tested occasionally on their reaction time to an extraneous stimulus (Keele & Boies, 1973; Stanners, Meunier, & Headley, 1969), monitored on a concurrent tracking task (Johnston, Greenberg, Fisher, & Martin, 1970), or compared on concurrent performances in sentence comprehension, sentence verification, or free recall of words (Baddeley & Hitch, 1974). The general finding was that the concurrent performances were worse when there was a simultaneous requirement of rehearsing a separate task. Further, the greater the memory load of the rehearsal task, the more detrimental the influence on the concurrent performances (Baddeley & Hitch, 1974; Johnston *et al.*, 1970). Stanners *et al.* (1969) also found that the later the scheduling of a reaction-time task into an unfilled rehearsal period, the faster was reaction time. As interpreted by Stanners *et al.*, longer periods of rehearsal prior to the reaction-time task allowed subjects to decrease their short-term memory loads by shifting memorial content to long-term storage. If so, a lessening of the rehearsal load may have freed capacities that were then reallocated to the reaction-time task.

In summary, the various evidences support the conclusion that rehearsal competes with other cognitive tasks for what appears to be a limited amount of processing time or capacity. Similarly, the outcomes are consistent with the inference that the structural capacities used in rehearsal are not exclusively dedicated to the rehearsal process.

C. TRACE STRUCTURE

Just as memorial representations in the short-term store do not necessarily carry the characteristics of the input modality, the traces that are regenerated and maintained in rehearsal do not necessarily carry all of the information available in the short-term trace (Atkinson & Shiffrin, 1968, pp. 92, 111). "Only that information selected by the subject, often a small proportion of the initial ensemble, is maintained."

Aside from qualitative changes introduced by rehearsers' selection strategies, traces that undergo rehearsal might be qualitatively altered by the process of rehearsal itself. If so, learners should be able to differentiate between memory traces formed during external presentations of stimuli and traces formed during rehearsal. Contrary to this expectation, the covert rehearsal of previously exposed pictures or words influenced estimates of the frequencies with which the pictures or words had actually been presented (Johnson, Raye, Wang, & Taylor, 1979; also see Shaughnessy & Underwood, 1973). Although this outcome suggests the two types of traces are qualitatively equivalent, it remains to be seen whether the frequency of external presentations of an item comparably influences estimates of rehearsal frequency.

If memorial traces are qualitatively altered by the process of rehearsal, it also should be possible to find variables that differentially affect rehearsed and unrehearsed traces. Empirical tests of this possibility have focused primarily on the variable of acoustic similarity. Short-term memory traces are known to be susceptible to interference from content that is acoustically similar (e.g., Conrad, 1964). Contrary to the hypothesis that memories formed during rehearsal would be less susceptible to acoustic interference, Sperling and Speelman (1970, p. 183) found that the effects of acoustic similarity were unrelated to rehearsal opportunity.

In another comparison of acoustic similarity on rehearsal effectiveness, Hayes and Rosner (1975) reported that preschoolers who were required to engage in overt, cumulative rehearsal had greater difficulty in remembering objects that had phonetically similar names. Conditions in which children labeled the objects at input, but did not engage in cumulative rehearsal, performed less well in recall but showed no difference in performance on the phonetically similar and dissimilar lists.

Disruptive effects from acoustic similarity also were reported by McFarland and Kellas (1974). During acquisition, learners allocated themselves longer periods of unfilled rehearsal time when lists were presented auditorily rather than visually. As interpreted by McFarland and Kellas, the acoustic traces from the most recently presented items interfered with the acoustic/articulatory representations of items already in the rehearsal buffer. When learners were told to rehearse cumulatively during the presentation, or to emphasize the initial portion of the list, the modality of presentation had no influence on acquisitional pacings.

The influence of acoustic similarity thus depends on the type of rehearsal used by learners. Rehearsal per se does not immunize the trace against the detrimental influences of acoustic similarity, but certain types of rehearsal result in traces that are less susceptible to interference from acoustic similarity. The existing data thus provide only a hedged answer to the question of whether a rehearsed trace is qualitatively identical to an unrehearsed trace.

In retrospect, it is unfortunate that the variable of acoustic similarity has provided the major test of the hypothesis that unrehearsed traces are qualitatively different from rehearsed traces. Rehearsers report that subvocal speech and thinking during covert rehearsal typically are accompanied by internalized auditory and articulatory stimuli. Thus, there is a priori reason for expecting that acoustic similarity would not differentiate the two types of traces.

Researchers, however, perhaps will uncover other variables that differentiate rehearsed from unrehearsed traces. Tzeng's (1976) research, for example, suggests that traces formed in rehearsal do not carry information regarding temporal order of occurrence. Learners typically have some temporal knowledge of the input order of items presented in a list (Underwood, 1977). However, if items are rehearsed together in combinations differing from the original input orders, different sets of temporal cues presumably would be formed. As a consequence, rehearsers seemingly would be less accurate in identifying the original input order. In fact, rehearsers who formed rehearsal sets that scrambled the original input orderings nevertheless were just as accurate in identifying the original input order as those who did not rehearse or those whose rehearsals were restricted to the currently presented word. Although Tzeng's finding suggests that rehearsal traces do not include temporal tags, another possibility is that temporal tags also were formed during rehearsal, but that rehearsers were able to keep such tags separate from the tags formed originally. Tzeng's evidence thus is not sufficient to allow the conclusion that the traces formed during rehearsal are qualitatively different from those formed during the original encoding.

V. Rehearsal as an Explanatory Variable: Directed Forgetting

Rehearsal has played a major theoretical role in explanations of various empirical findings. Theoretical accounts of directed forgetting, for example, frequently have pitted a rehearsal interpretation against an explanation emphasizing more efficient search operations during retrieval. Through an examination of the directed forgetting literature, the present section illustrates the usage of rehearsal in explaining empirical outcomes in a particular content area.

Within the directed forgetting paradigm, items are designated either as content to be remembered (*TBR*) or forgotten (*TBF*). The designation of "remember" or "forget" has been provided as early as prior to the presentation of any subunits (Geiselman, 1975), simultaneously with presentation (e.g., Roediger & Crowder, 1972), from 0 to 12 sec after the presentation of each subunit or groupings of subunits (e.g., Woodward *et al.*, 1973), or as late as immediately prior to the test of retention (Epstein, 1969). Although some investigators have attempted to minimize selective rehearsal of the *TBR* content, learners typically have had some opportunity for selective rehearsal either during the presentation of list items (Bruce & Papay, 1970; Davis & Okada, 1971; Geiselman, 1974, 1975; Gorfein, Arbak, Phillips, & Squillace, 1976; Reitman, Malin, Bjork, & Higman, 1973; Timmins, 1973; Woodward *et al.*, 1973; Woodward, Park, & Seebohn, 1974), during separate presentations of remember ("*R*") or forget ("*F*") cues (Bjork, 1970b; Bjork & Geiselman, 1978; MacLeod, 1975), or during retention intervals following receipt of the "*R*" and "*F*" cues (Epstein, Massaro, & Wilder, 1972; Epstein & Wilder, 1972; Geiselman, 1974, 1975; Johansson, 1975; Jongeward *et al.*, 1975; Reed, 1970; Roediger & Crowder, 1970; Shebilske & Epstein, 1973; Shebilske, Wilder, & Epstein, 1971; Timmins, 1973). Regardless of whether learners are given advance warning that the remembering of "forget" items also will be tested, the "remember" items are much more likely to be recalled. Tests of delayed recognition sometimes show equivalent accuracies in identifying *TBR* and *TBF* items (Block, 1971; Elmes, Adams, & Roediger, 1970), but a substantial number of experimenters have reported superior accuracies in recognizing *TBR* units (Bjork & Geiselman, 1978, Exp. I; Bruce & Papay, 1970; Davis & Okada, 1971;

MacLeod, 1975; Roediger & Crowder, 1972; Woodward *et al.*, 1973, 1974). Consistent with a selective rehearsal interpretation, the superiority evidenced in remembering "*R*" items is offset by an equivalent inferiority in remembering "*F*" items (Geiselman, 1974; Reitman *et al.*, 1973). Although such outcomes have been labeled "directed forgetting," there are no serious advocates of the position that learners simply dump or erase the "*F*" items from memory. Instead, the directed forgetting cues are presumed to trigger a strategy of selective rehearsal of the *TBR* items. Important components of an effective rehearsal strategy apparently include the implicit or explicit retrieval of *TBR* items and nonretrieval of *TBF* items at the time the items are designated as "*R*" or "*F*" (Bjork & Geiselman, 1978) as well as other processes that allow differentiation of the *TBR* and *TBF* units into separate subsets (Bjork, 1972; Reitman *et al.*, 1973). "Directed forgetting" thus appears to be an inappropriate label for what is actually "directed learning."

Aside from the differential remembering of *TBR* and *TBF* subsets, learners who receive a postpresentation cue informing them to remember only one of the two subsets of an input list perform better on the designated subset of items than learners directed to recall first the designated subset and then the remaining subset (e.g., Shebilske *et al.*, 1971). This outcome, termed the "only" effect, has been attributed to reduced interference from the *TBF* subset through cuing that allows learners to restrict their retrieval search to items in the *TBR* subset (Shebilske *et al.*, 1971; Epstein *et al.*, 1972). For unexplained reasons, the "only" effect is more likely to be evident when the second input subset is tested (e.g., Block, 1971; Epstein, 1969, 1970), but some experimenters also have found an "only" effect in tests of the first subset (e.g., Shebilske *et al.*, 1971). Consistent with the selective search interpretation, the "only" effect was larger when the two subsets constituted distinctive categories (Shebilske *et al.*, 1971), and the effect was eliminated when the differentiation into two subset categories was blurred either by a random or regular interspersing of the two types of subset items during presentation (Epstein, 1969; Woodward *et al.*, 1974). Contrarily, however, the effect is larger when both input subcategories are presented in the same sensory modality, e.g., visual, rather than different modalities, e.g., visual and auditory (Epstein, 1970). The "only" effect also disappeared on a matching task in which the response alternatives were restricted to the "remember" subset, but the effect was in evidence when the matching task included responses from both subsets (Epstein *et al.*, 1972). As interpreted from the selective search viewpoint, the absence of the "only" effect when matching was restricted to the *TBR* subset occurred because the size and composition of the appropriate search set already was delimited by the available alternatives. The search set presumably was equivalent regardless of whether the learner was told to be responsible for one or both subsets, whereas a directed forgetting cue supposedly could aid search when all of the alternatives were potential matches.

The "only" effect and the differential memory for "*R*" and "*F*" items both are a consequence of cues to forget, but the two phenomena thus have received different explanations (Jongeward *et al.*, 1975; Woodward *et al.*, 1974). Differences in the remembering of "*R*" and "*F*" units have been attributed to differential encoding engendered by selective rehearsal, whereas the superior remembering of "*R*" units in the "only" effect has been interpreted as an output phenomenon caused by selective search operations during retrieval or by differential short-term memory loads. Experimentally, however, the methodological sequence that produces the "only" effect and the differential remembering of "*R*" and "*F*" items is exactly the same. The only difference occurs in experimenters' selection of a baseline against which to compare the remembering of the "*R*" subunits—either the remembering of (*a*) "*F*" units, or else (*b*) "*R*" units when both subsets are designated as required learning.

Despite the similarity in the procedural operations that produce the "only" effect and the differential remembering of "*R*" and "*F*" items, the empirical advantage of the directed forgetting group in the two types of comparison may stem from different blendings of selective rehearsal and selective search. If so, critical determinants of the causative blend would appear to be the time at which the "*F*" cue is received by the learner and the nature of the cognitive processing that follows the "*F*" cue. When an "*F*" cue is received immediately prior to the test of retention, the learner can omit "*F*" items in output, and may show enhanced performance on the "*R*" subset, but there is no evidence of a permanent erasure (Jongeward *et al.*, 1975). When the "*F*" cues are received sufficiently early to direct rehearsal, the rehearser can allocate processing efforts to the learning of "*R*" items. The mere occurrence of longer rehearsal times, however, does not always result in larger differentials in remembering "*R*" and "*F*" items. Although the relative advantage of "*R*" items is somewhat greater when cuing is provided earlier in the rehearsal period (Jongeward *et al.*, 1975; Reed, 1970), the differential remembering of the "*R*" and "*F*" items depends upon an "*F*" cue followed by rehearsal time rather than rehearsal time prior to the receipt of the "*F*" cue (Davis & Okada, 1971; Woodward *et al.*, 1973). In the latter two experiments, the subjects apparently engaged in a maintenance holding-type rehearsal rather than attempting to encode the materials into long-term memory.

To summarize, early receipt of the "*F*" cue increases the likelihood of selective rehearsal. When the "*F*" cue is received immediately prior to recall, and the learner has no prior basis for inferring which category will be designated the *TBR* subset, selective rehearsal is minimized, and a difference in remembering is more likely to represent an output phenomenon. Even so, the issue is not closed, and selective search may play only a minor causative role in the "only" effect. Given the necessity to assess the independent influence of selective search, it is unfortunate that most experimenters simply conceded the likelihood that selective search and selective rehearsal both are determinants of the superior remembering of *TBR* items (e.g., Jongeward *et al.*, 1975). As a consequence, perhaps, most investigators of the "only" effect have provided "*F*" cues at the beginning of the retention interval rather than immediately prior to the test of retention (Bjork, 1970b; Epstein & Wilder, 1972; Epstein *et al.*, 1972; Johansson, 1975; Jongeward *et al.*, 1975; Reed, 1970; Shebilske & Epstein, 1973; Shebilske *et al.*, 1971). Although some experimenters have required distractor activities such as backward counting during the brief retention intervals, the number of units in the *TBR* subset has not exceeded four, and the learners may have engaged in some maintenance rehearsal.

Perhaps the most convincing demonstration of the "only" effect in the absence of rehearsal was Block's (1971) Experiment II. Although the "*F*" cue to forget the first six words was presented simultaneously with the presentation of the second *TBR* subset of six words, rehearsal opportunity was lessened by a rapid presentation rate of .6 sec per word and the absence of a retention interval prior to recall. An "only" effect was evident in immediate recall, but delayed recognition performances were unrelated to the presence or absence of an "*F*" cue.

Only Epstein (1969, 1970) and Bartz (1972) delayed the introduction of the "*F*" cue until the beginning of the recall period. Bartz found an "only" effect only when the subjects, like Epstein's subjects, were not given advance warning regarding the output order that would be required when both subsets were to be recalled. According to Bartz, when subjects had advance knowledge of output order, and therefore knew that a particular subset would be either the first or only output on three-fourths of the trials, the subjects thereby were more limited in the range of rehearsal strategies that could be used during presentation.

The selective search view thus needs elaboration to account for (*a*) the disappearance of the effect when subjects possess advance knowledge of the order of recalling two subsets, (*b*) the more reliable occurrence of the effect when the second input category is tested (Epstein, 1969), and (*c*) the larger effect when the subset categories are presented in the same sensory mode (Epstein, 1970). A more convincing explanation also is needed for how an "*F*" cue can facilitate matching performances when all alternatives of both subsets are already available to the learner (Epstein *et al.*, 1972). Since retrieval presumably was not a problem in the matching task, and the two subsets of three paired associates already were potentially differentiable on the basis of temporal (first vs second subset) and categorical membership (parts of houses vs insects), it is not obvious how an "*F*" cue would allow a more effective retrieval search.

Selective rehearsal, in turn, is unlikely to be toppled as an alternative explanation as long as experimenters permit subjects opportunity for rehearsal. Since selective rehearsal is likely to be enhanced by manipulations that result in increased differentiation of the two subsets (Bjork, 1970b, 1972), the rehearsal viewpoint can readily explain the same findings that can be interpreted by a selective search theory. To account for the facilitative effects of an "*F*" cue on later matching performances for "*R*" items (Epstein *et al.*, 1972), a rehearsal theorist might capitalize on the existence of opportunity for selective rehearsal. To explain Epstein's (1969, 1970) finding that the "only" effect is more likely when the second input category is tested, a rehearsal theorist might speculate that learners were engaged primarily in a maintenance rehearsal in which later input displaced the earlier input. The larger effect when the input categories are presented in the same sensory mode (Epstein, 1970), however, presents an interpretative difficulty for the selective rehearsal viewpoint as well as for the selective search interpretation.

In summary, selective rehearsal provides a viable account of the superior remembering of "*R*" units over "*F*" units. With respect to the superior remembering of "*R*" units in the "only" effect, the source of the superiority is uncertain, and may be a consequence of more effective search operations during retrieval, selective rehearsal prior to retrieval, or a combination of selective search and selective rehearsal.

VI. Theoretical Models of Rehearsal

Postrehearsal performances presumably are a function of new structural representations formed during rehearsal. Does the process of rehearsal thus lead to the formation of separate residual traces? If so, does each rehearsal result in another internalized copy of the item? Alternately, does rehearsal augment the memorial representation formed during the original encoding? What functions are served by rehearsal? What are the processes by which such functions are achieved? What are the variables that influence rehearsal? Although such questions are fundamental in understanding learning and memory, there has been virtually no theorizing on such issues.

A number of researchers have advanced one or more hypotheses about rehearsal, but only Atkinson and Shiffrin (1968) and Bernbach (1969) have attempted a systematic account of rehearsal functioning. Though more than a decade has passed, the two theories have received few direct empirical tests. As will become evident, however, there are grounds for suspecting that a number of their theoretical assumptions may be incorrect, and that new theoretical models are needed.

A. BERNBACH'S REPLICA MODEL

When an item is presented, according to Bernbach (1969, 1970), an internal representation of the item, called a replica, is formed in memory. During the time remaining in the presentation interval, the learner was assumed to rehearse either the item itself or an item presented earlier. The selection of previously presented items for additional rehearsal was presumed to be randomly determined. Each rehearsal was assumed to result in the formation and storage of an additional replica of the item rehearsed. Remembering, however, did not depend directly upon the number of replicas. As long as one replica of the item remained in storage, it was assumed that a correct response would be made. The number of replicas did influence memory, however, in increasing the probability that at least one replica would remain in storage after interference has taken its toll.

Bernbach's theory, though admittedly based upon simplifying assumptions, provides an excellent mathematical description or prediction of a wide variety of experimental findings across a variety of experimental paradigms. The theory offers an account of primacy and recency effects, the shapes of forgetting curves, the influence of intervening items, and other well-known outcomes in verbal learning.

The success of Bernbach's theory, however, cannot be attributed to well-grounded assumptions regarding rehearsal. First, the assumption that multiple replicas are formed by rehearsal has not been substantiated, and it is not obvious as to the type of evidence that would allow the inference that multiple traces had been formed. The occurrence of sigmoid-shaped curves in short-term forgetting was interpreted by Bernbach as evidence supporting the predictions of the multitrace model (see also Pollatsek & Bettencourt, 1976). However, slight modifications in the assumptions of trace-strength or multistore models also could produce satisfactory curve fitting (Bjork, 1970a). Success in predicting such data thus does not provide convincing evidence that rehearsal increases the number of trace representations.

In justifying the assumption that rehearsers randomly select items for further rehearsal, Bernbach (1969, p. 209) concedes that subjects might engage in more systematic rehearsal strategies. However, the choice was defended on the basis that individual subjects presumably pursue different rehearsal strategies, and that the intent of the model was to account only for group performances. Since Bernbach's decision was made, however, a considerable body of evidence has accumulated indicating that rehearsers as a group do show systematic biases in selecting items for additional rehearsal (e.g., Einstein *et al.*, 1974; Johnson, 1975; Rundus, 1974).

Empirical data also are needed to support Bernbach's assumptions regarding the speed of rehearsing and the allocation of rehearsals. During 2-sec presentation intervals, rehearsers were assumed to encode the item, and then engage in 2.65 rehearsals of the item plus an additional 2.83 rehearsals of previously presented items. The rationale for such rehearsal allocation, however, was based upon statistical rather than psychological considerations. Aside from the arbitrariness of the rehearsal allocations, the resultant rate of 309 msec for each replica/rehearsal appears excessively rapid, especially when switching time is needed for transitions among the three types of activities.

B. ATKINSON AND SHIFFRIN/THE REHEARSAL BUFFER

More than 10 years after its publication, Atkinson and Shiffrin's (1968) theory of human memory remains the most important theoretical account of rehearsal processes. Consistent with their belief that "rehearsal is one of the most important factors in experiments in human memory," their theoretical centerpiece was a "rehearsal buffer" that mediated events controlling memory. Rehearsal was conceptualized narrowly as a repetitive recycling that regenerated traces within the buffer. Longer times in the buffer, however, were assumed to increase the probability of direct entry into long-term storage and to increase the opportunity for coding and other storage processes. Without rehearsal, in contrast, items in the buffer were assumed to be ousted by the entry of new items or lost through decay.

1. Buffer Capacity

Figure 1 depicts Atkinson and Shiffrin's (1968) conception of the structural relationships between the rehearsal buffer and other components of the memory system. Aside from the supposition that rehearsal regenerations carry less information than the short-term traces, the short-term store and the rehearsal buffer are differentiated conceptually by larger

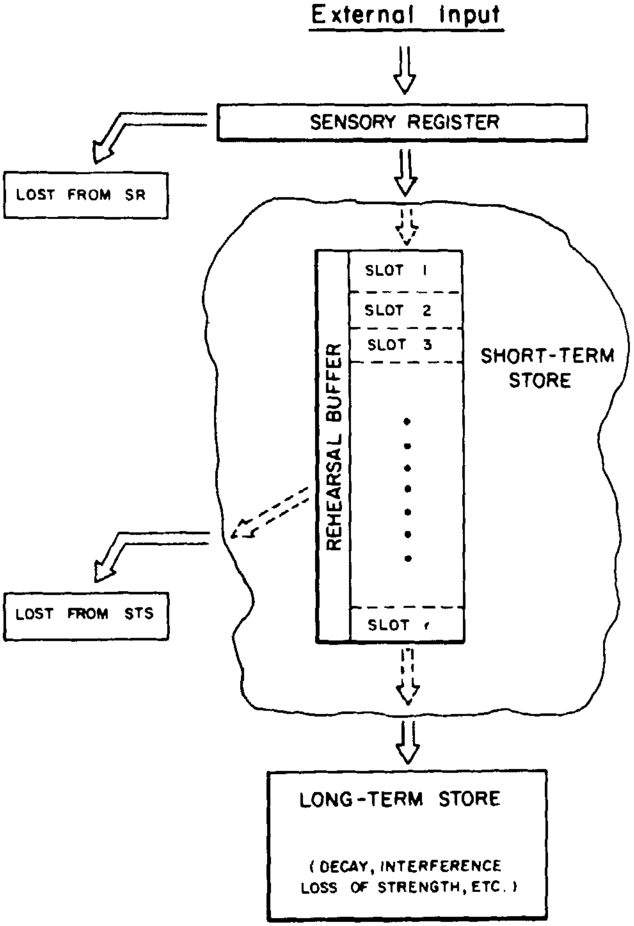


Fig. 1. Atkinson and Shiffrin's (1968) rehearsal buffer and its relationship to other components of the memory system.

capacity in the short-term store. Operationally, however, rehearsal capacity is estimated from the traditional digit-span measure of short-term memory capacity. Even with an alternative measure, precise estimates of rehearsal capacity would require insight into the size of the units actually being rehearsed (Atkinson & Shiffrin, 1968, pp. 127, 166–167), the rate of decay in short-term storage (p. 112), the capacity usurped by testing events (p. 131), the time devoted to search operations (p. 131), and the extent to which rehearsers engage in coding operations rather than repetitive rehearsal (pp. 114–115, 187–190).

2. *Entry into Rehearsal Buffer*

Atkinson and Shiffrin (p. 113) assumed that "the maintenance and use of the buffer is a process entirely under the control of the subject." Rehearsers thus can elect not to enter particular items into the buffer (p. 114). Rehearsers also have the option of not forming a buffer (p. 183) or decreasing the size of the buffer to as small as a single item (pp. 187–188). The preceding assumptions, like others that follow, are not derived from empirical evidence. Unless indicated otherwise, the postulates regarding rehearsal functioning are neither supported nor refuted by empirical data.

Though rehearsal decisions are made individually, rehearsers tend to develop common strategies based upon previous rehearsal experiences. Thus, there is predictability to rehearsal patterns. Faster rates of presentation presumedly increase the probability that a buffer will be formed (p. 182), whereas slower rates presumedly increase the likelihood that long-term encodings will be attempted (p. 187). The formation of rehearsal buffers also was asserted to be more likely when learners assume that short-term rememberings will suffice (p. 116). With memory loads equal to or less than the maximal size of the buffer, repetitive rehearsal activities presumably are encouraged (p. 166). When memory loads exceed the buffer, however, rehearsers presumably are more likely to form rehearsal sets of smaller size and then spend the remaining time in long-term coding activities.

Entry of new items into the buffer presumably can be impeded by the presence of items already in the buffer. As suggested by Atkinson and Shiffrin (p. 114), "some combinations of items are particularly easy to rehearse, making the subject loath to break up the combination. In fact, the work involved in introducing a new item into the buffer and deleting an old one may alone give the subject incentive to keep the buffer unchanged."

Once items are well known, however, rehearsers supposedly are reluctant to reenter well-known items into the buffer on succeeding trials. Unlike the entry suppositions noted so far, empirical evidence was cited for this assumption (Atkinson and Shiffrin, 1968, pp. 158–159). In an experiment by W. Thomson, two paired associates were presented on several different trials and tested at various lags. Other paired associates received new response pairings after each test trial. With longer lags, recall declined for pairings presented only once—a fact attributed to such items having been knocked out of the buffer by the entry of other items. Since recall of the well-known items showed little decline with longer lag lengths, Atkinson and Shiffrin assumed that these items had not been in the buffer. A simpler view, however, might be that the superior recall of the well-known items was a consequence of superior associative strength. Furthermore, other research indicates that rehearsers show some tendency to rehearse associations that are already known (Ciccone & Brelsford, 1974; Johnson, 1975).

Entry into the buffer also is more likely when learners select some subset of the items for special attention or rehearsal (Atkinson & Shiffrin, 1968, p. 116). In support of this assertion, Atkinson and Shiffrin cite Harley's (1965a,b) evidence that subjects recalled best those items that led to monetary rewards. Indirect confirmatory evidence also may be found now in studies suggesting that learners allocate more rehearsals to "remember" items than to "forget" items (e.g., Jongeward *et al.*, 1975). The primacy effect, similarly is attributed to extra rehearsal on the initial items of a list (Atkinson & Shiffrin, 1968, pp. 174–183). Recency effects, in contrast, are attributed to rehearsers' ability to recall accurately the existing contents of the rehearsal buffer. Congruent with expectations, the interpolation of an intervening task eliminated the recency effect, presumably because the buffer had been cleared, whereas the primacy effect survived (Postman & Phillips, 1965).

3. *Buffer Operations*

For items in the buffer, rehearsers appear to have knowledge of the temporal order of entry (Atkinson & Shiffrin, 1968, pp. 162–163). In a study by Freund and Rundus, paired associates were presented once and then tested later in the presentation sequence. When learners could supply the correct response, they also had accurate knowledge of the lag between presentation and the test trial. When the lag length exceeded five, however, or when learners could not supply the correct response to the stimulus probe, the lag judgments were virtually unrelated to the true lag.

Though rehearsers have knowledge of the order in which items are entered into the rehearsal buffer, the ordering within the buffer itself does not necessarily correspond to the entry order (Atkinson & Shiffrin, 1968, p. 127). Rehearsers thus presumably can rearrange the contents of the buffer. Further, rehearsers apparently also have the option of entering items from long-term memory into the short-term rehearsal buffer slots.

Yet, some entry arrangements may facilitate learning more than others. Imhoff, Horton, Weldon, and Phillips (1977) varied the number of items presented simultaneously while holding constant the list length and average study time per item. Presentation groupings of one, three, or six items were followed by unfilled rehearsal periods of .5 sec per item. As the size of the input grouping increased, final recall decreased. Imhoff *et al.* (1977) suggest that the larger set sizes forced learners into allocating some processing capacity to maintenance activity rather than long-term encoding. Usage of the buffer thus may be more effective when rehearsal efforts can be devoted exclusively to a single item prior to the arrival of a second item. The conclusion, however, must be regarded as tentative since the methodology of Imhoff *et al.* confounds the simultaneity variable with the number and duration of scheduled rehearsal periods. Under other conditions of learning, such as concept formation, simultaneous presentation of items is more effective than consecutive presentation (e.g., Crouse & Duncan, 1963), and the advantage has been attributed to the lessening of memory loads. The question thus remains open as to whether contiguous entry into the buffer promotes the establishment of associations between such items more than the entry of target pairs separated by one or more intervening items.

When a previously encountered item is presented again, and the item already occupies a slot in the rehearsal buffer, the unit in residence is presumed to be automatically replaced by an updated trace that is less likely to be deleted. Since old items, by definition, include paired associates in which previously presented stimuli now are paired with new responses, several counterintuitive predictions are generated. In contrast with expectations generated from interference theory or from knowledge regarding negative transfer, the Atkinson and Shiffrin (1968, pp. 147–148) model can explain Brelsford, Shiffrin, and Atkinson's (1968) finding that the remembering of a new pairing improved as a function of the number of previous pairings of a stimulus with other responses. Similarly, Atkinson and Shiffrin (1968, pp. 149–151) easily account for the finding of Brelsford *et al.* (1968) that shorter lags between the two presentations resulted in superior recall of the new response. Shorter lags presumably increased the probability that the original paired associate was in the buffer when the new pairing was presented. The second pairing thus would be automatically entered into the buffer and be available for recall. Interference theory, in contrast, might have predicted that longer lags would be beneficial since this would allow additional opportunity for forgetting the original interfering association.

Atkinson and Shiffrin's (1968) view that initial pairings are replaced by new pairings, however, appears undermined by Bruce and Weaver's (1973) finding that the short-term recall of an A–B pair was *augmented* by the subsequent appearance of a related A–D pair. Beyond this, there is need for detailing the process whereby a stimulus with a new response is somehow recognized and then automatically entered into the buffer as a replacement. Why, for example, is there not equal priority for an "old" response that now is paired with a "new" stimulus? Further, does recognition and replacement fail when the two stimuli are similar but not identical? Is the presentation of each item accompanied by a scan of the buffer to determine whether a match can be made? Can learners also recall the first response? If so, such recall would indicate that the initial paired associate was not automatically replaced by the second pairing.

4. *Deletion from Buffer*

Atkinson and Shiffrin (1971, p. 85) indicate that "what item is replaced depends upon a number of psychological factors, but in the model the decision is approximated by a random process." With covert study procedures, the assumption of random replacement allowed accurate prediction of the relationship between recall and test lag. When learners were required to recite each item aloud during presentation, however, the retention curve showed an S-shaped decline that did not fit the predictions of the model. Thereupon, Atkinson and Shiffrin (1968, p. 122) assumed that the learners in the overt condition were deleting the oldest items first rather than deleting items on a random basis. No psychological rationale was offered for why the deletion rules differed in the two study procedures. Like the earlier assumption that an old stimulus with a new response qualified as a "previously encountered" item, the differing deletion assumptions are based upon mathematical curve-fitting requirements rather than an explicit psychological rationale.

C. THE REHEARSAL BUFFER—A DECADE LATER

Atkinson and Shiffrin's (1968) theoretical formulations have received only one serious empirical challenge. Alas, as luck would have it, the postulate that received the most experimental attention, the assertion that the length of stay in the buffer is directly related to the probability of long-term storage, was not substantiated. Postrehearsal gain typically is positively related to time spent in rehearsing (see Section III,A), but longer periods of maintenance rehearsal do not necessarily increase the probability of long-term recall (e.g., Craik & Watkins, 1973; Jacoby & Bartz, 1972; Meunier *et al.*, 1972; Rundus, 1977; Woodward *et al.*, 1973). Given the meager amount of post-1968 data relevant to the notion of a rehearsal buffer, it is not obvious that a competing theoretical system could be developed now that would represent a quantum advance over Atkinson and Shiffrin's formulations. Yet, clearly there is need for extending their efforts. Aside from testing their empirical assumptions, there also is need for reexamining their theoretical tenets. Perhaps the most basic issue is whether the construct of a rehearsal buffer is still the most useful way of conceptualizing the rehearsal process. Although the construct provides heuristic advantages in conceptualizing a flow pattern of events, the buffer notion could be replaced with the simpler conception that short-term memory traces lose strength unless renewed by additional presentations or by rehearsal. Since the buffer has been conceptualized as a limited-capacity workspace and storage, advocates of the buffer concept seemingly must demonstrate that items are ejected from the buffer by the entrance of new items. Similarly, the limited-space hypothesis allows the corollary that items cannot be entered into a full buffer unless new space is allocated through the ejection of old items. As noted, some empirical evidence suggests that items are not ousted from the buffer by the entrance of replacement items (e.g., Bruce & Weaver, 1973).

Regardless of the fate of the buffer, there is need to test the generality of Atkinson and Shiffrin's theoretical assumptions. In developing their theory, Atkinson and Shiffrin elected to restrict the explanatory domain to rote rehearsal. Through the use of a task in which rehearsers were "required to keep track of constantly changing responses associated with a fixed set of stimuli," Atkinson and Shiffrin (1968, pp. 123–124) thereby convinced the rehearser that "long-term storage operations, such as coding, are not particularly useful." While such restrictions are strategically useful in answering certain questions regarding buffer functioning, it seems likely that rehearsers would use their rehearsal capabilities differently in attempting to encode content meaningfully in preparation for a test of long-term memory.

Theorists of the second decade hopefully will attempt to account for rehearsers' performances under experimental conditions that are less restrictive. Given the situation in which rehearsers are aware that long-term recall will be assessed, is there an increase in the number of rehearsals? Is the rehearser more likely to stagger the rehearsals of particular items? Are longer rehearsal times spent on each item? Is the rehearser more likely to engage in coding that involves rehearsing several items together? Will rehearsers more actively engage in attempts to relate the content of rehearsal to existing semantic knowledge? What particular coding strategies are used by rehearsers? How effective are such techniques? What are the psychological biases that guide rehearsers in selecting the content that will receive further rehearsal? Clearly there is ample opportunity for empirical and theoretical questioning of Atkinson and Shiffrin's views of the rehearsal process.

REFERENCES

Ashcraft, M. H., Kellas, G., & Needham, S. Rehearsal and retrieval processes in free recall of categorized lists *Memory and Cognition,* 1975, **3,** 506–512.

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2). New York: Academic Press, 1968.

Atkinson, R. C., & Shiffrin, R. M. The control of short-term memory. *Scientific American*, 1971, **225**, 82-90.

Baddeley, A. D. The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review*, 1978, **85**, 139-152.

Baddeley, A. D., & Hitch, G. Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8). New York: Academic Press, 1974.

Bandura, A., & Jeffery, R. W. Role of symbolic coding and rehearsal processes in observational learning. *Journal of Personality and Social Psychology*, 1973, **26**, 122-130.

Bandura, A., Jeffery, R., & Bachicha, D. L. Analysis of memory codes and cumulative rehearsal in observational learning. *Journal of Research in Personality*, 1974, **7**, 295-305.

Bartz, W. H. Rehearsal strategies and partial recall in immediate memory. *Journal of Experimental Psychology*, 1972, **94**, 141-145.

Bellezza, F. S., & Cheney, T. L. Isolation effect in immediate and delayed recall. *Journal of Experimental Psychology*, 1973, **99**, 55-60.

Bellezza, F. S., Geiselman, R. E., & Aronovsky, L. A. Eye movements under different rehearsal strategies. *Journal of Experimental Psychology: Human Learning and Memory*, 1975, **1**, 673-679.

Bellezza, F. S., & Walker, R. J. Storage-coding trade-off in short-term store. *Journal of Experimental Psychology*, 1974, **102**, 629-633.

Belmont, J. M., & Butterfield, E. C. What the development of short-term memory is. *Human Development*, 1971, **14**, 236-248.

Bernbach, H. A. Replication processes in human memory and learning. In J. T. Spence & G. H. Bower (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 3). New York: Academic Press, 1969.

Bernbach, H. A. A multiple-copy model for postperceptual memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

Bjork, R. A. Repetition and rehearsal mechanisms in models for short-term memory. In D. A. Norman, (Ed.), *Models of human memory*. New York: Academic Press, 1970. (a)

Bjork, R. A. Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 255-268. (b)

Bjork, R. A. Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972.

Bjork, R. A., & Geiselman, R. E. Constituent processes in the differentiation of items in memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, **4**, 347-361.

Bjork, R. A., & Woodward, A. E. Directed forgetting of individual words in free recall. *Journal of Experimental Psychology*, 1973, **99**, 22-37.

Block, R. A. Effects of instructions to forget in short-term memory. *Journal of Experimental Psychology*, 1971, **89**, 1-9.

Bower, G. H., & Reitman, J. S. Mnemonic elaboration in multilist learning. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 478-485.

Brelsford, J. W., Jr., Shiffrin, R. M., & Atkinson, R. C. Multiple reinforcement effects in short-term memory. *British Journal of Mathematical and Statistical Psychology*, 1968, **21**, 1-19.

Broadbent, D. E. *Perception and communication*. New York: Pergamon, 1958.

Brodie, D. A., & Prytulak, L. S. Free recall curves: Nothing but rehearsing some items more or recalling them sooner? *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 549-563.

Brown, J. Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 1958, **10**, 12-21.

Bruce, D., & Papay, J. P. Primacy effect in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 473-486.

Bruce, D., & Weaver, G. E. Retroactive facilitation in short-term retention of minimally learned paired associates. *Journal of Experimental Psychology,* 1973, **100,** 9-17.

Buschke, H., & Hinrichs, J. V. Controlled rehearsal and recall order in serial list retention. *Journal of Experimental Psychology,* 1968, **78,** 502-509.

Buschke, H., & Kintsch, W. Rehearsal strategies and the serial-position curve in immediate free recall of ordered items. *Quarterly Journal of Experimental Psychology,* 1970, **22,** 347-352.

Butterfield, E. C., & Belmont, J. M. Relations of storage and retrieval strategies as short-term memory processes. *Journal of Experimental Psychology,* 1971, **89,** 319-328.

Carter, J. F., & Van Matre, N. H. Note taking versus note having. *Journal of Educational Psychology,* 1975, **67,** 900-904.

Cermak, L. S. Rehearsal strategy as a function of recall expectation. *Quarterly Journal of Experimental Psychology,* 1972, **24,** 378-385.

Cermak, L. S., & Craik, F. I. M. (Eds.), *Levels of processing and human memory.* Hillsdale, N.J.: Erlbaum, 1978.

Cermak, L. S., & Levine, R. Encoding as a function of the presentation-rehearsal interval in short-term memory. *Psychonomic Science,* 1971, **23,** 423-424.

Ciccone, D. S., & Brelsford, J. W. Interpresentation lag and rehearsal mode in recognition memory. *Journal of Experimental Psychology,* 1974, **103,** 900-906.

Cohen, G. How are pictures registered in memory? *Quarterly Journal of Experimental Psychology,* 1973, **25,** 557-564.

Cohen, R. L. The differential effect of stimulus intensity on rehearsed and unrehearsed material in short-term auditory memory. *Canadian Journal of Psychology,* 1967, **21,** 277-284.

Colegate, R. L., & Eriksen, C. W. Implicit speech as an encoding mechanism in visual perception. *American Journal of Psychology,* 1970, **83,** 208-215.

Conrad, R. Acoustic confusions in immediate memory. *British Journal of Psychology,* 1964, **55,** 75-84.

Craik, F. I. M., & Watkins, M. J. The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 599-607.

Crouse, J. H., & Duncan, C. P. Verbal concept sorting as a function of response dominance and sorting method. *Journal of Verbal Learning and Verbal Behavior,* 1963, **2,** 480-484.

Cuvo, A. J. Incentive level influence on overt rehearsal and free recall as a function of age. *Journal of Experimental Child Psychology,* 1974, **18,** 167-181.

Dark, V. J., & Loftus, G. R. The role of rehearsal in long-term memory performance. *Journal of Verbal Learning and Verbal Behavior,* 1976, **15,** 479-490.

Darley, C. F., & Glass, A. L. Effects of rehearsal and serial list position on recall. *Journal of Experimental Psychology: Human Learning and Memory,* 1975, **1,** 453-458.

Davis, J. C., & Okada, R. Recognition and recall of positively forgotten items. *Journal of Experimental Psychology,* 1971, **89,** 181-186.

DeRemer, P., & D'Agostino, P. R. Locus of distributed lag effect in free recall. *Journal of Verbal Learning and Verbal Behavior,* 1974, **13,** 167-171.

Detterman, D. K. The von Restorff effect and induced amnesia: Production by manipulation of sound intensity. *Journal of Experimental Psychology: Human Learning and Memory,* 1975, **1,** 614-628.

Detterman, D. K., & Ellis, N. R. Determinants of induced amnesia in short-term memory. *Journal of Experimental Psychology,* 1972, **95,** 308-316.

Dillon, R. F., & Reed, L. S. Short-term memory as a function of information processing during the retention interval. *Journal of Experimental Psychology,* 1969, **81,** 261-269.

Einstein, G. O., Pellegrino, J. W., Mondani, M. S., & Battig, W. F. Free-recall performance as a function of overt rehearsal frequency. *Journal of Experimental Psychology,* 1974, **103,** 440–449.

Ellis, J. A., & Montague, W. E. Effect of recalling on proactive interference in short-term memory. *Journal of Experimental Psychology,* 1973, **99,** 356–359.

Ellis, N. R. Memory processes in retardates and normals. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 4). New York: Academic Press, 1970.

Elmes, D. G., Adams, C., III, & Roediger, H. L., III. Cued forgetting in short-term memory. *Journal of Experimental Psychology,* 1970, **86,** 103–107.

Elmes, D. G., Sanders, L. W., & Dovel, J. C. Isolation of massed- and distributed-practice items. *Memory and Cognition,* 1973, **1,** 77–79.

English, H. B., & English, A. C. *A comprehensive dictionary of psychological and psychoanalytical terms.* New York: Longmans, Green, 1958.

Epstein, W. Poststimulus output specification and differential retrieval from short-term memory. *Journal of Experimental Psychology,* 1969, **82,** 168–174.

Epstein, W. Facilitation of retrieval resulting from post-input exclusion of part of the input. *Journal of Experimental Psychology,* 1970, **86,** 190–195.

Epstein, W., Massaro, D. W., & Wilder, L. Selective search in directed forgetting. *Journal of Experimental Psychology,* 1972, **94,** 18–24.

Epstein, W., & Wilder, L. Searching for to-be-forgotten material in a directed forgetting task. *Journal of Experimental Psychology,* 1972, **95,** 349–357.

Estes, W. K. Phonemic coding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 360–372.

Ferguson, R. P., & Bray, N. W. Component processes of an overt rehearsal strategy in young children. *Journal of Experimental Child Psychology,* 1976, **21,** 490–506.

Fisher, M. A., & Zeaman, D. An attention-retention theory of retardate discrimination learning. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 6). New York: Academic Press, 1973.

Flexser, A. J. Long-term recognition latencies under rehearsal-controlled conditions: Do list-length effects depend on active memory? *Journal of Experimental Psychology: Human Learning and Memory,* 1978, **4,** 47–54.

Frank, H. S., & Rabinovitch, M. S. Auditory short-term memory: Developmental changes in rehearsal. *Child Development,* 1974, **45,** 397–407.

Fritzen, J. Intralist repetition effects in free recall. *Journal of Experimental Psychology: Human Learning and Memory,* 1975, **1,** 756–763.

Geiselman, R. E. Positive forgetting of sentence material. *Memory and Cognition,* 1974, **2,** 677–682.

Geiselman, R. E. Semantic positive forgetting: Another cocktail party problem. *Journal of Verbal Learning and Verbal Behavior,* 1975, **14,** 73–81.

Glenberg, A., Smith, S. M., & Green, C. Type I rehearsal: Maintenance and more. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 339–352.

Gorfein, D. S., Arbak, C., Phillips, R. V., & Squillace, L. Within-list serial position effects in free recall. *Journal of Experimental Psychology: Human Learning and Memory,* 1976, **2,** 161–171.

Götz, A., & Jacoby, L. J. Encoding and retrieval processes in long-term retention. *Journal of Experimental Psychology,* 1974, **102,** 291–297.

Harley, W. F., Jr. The effect of monetary incentive in paired associate learning using a differential method. *Psychonomic Science,* 1965, **2,** 377–378. (a)

Harley, W. F., Jr. The effect of monetary incentive in paired associate learning using an absolute method. *Psychononic Science,* 1965, **3,** 141–142. (b)

Hayes, D. S., & Rosner, S. R. The phonetic effect in preschool children: The influence of overt rehearsal and verbal instructions. *Journal of Experimental Child Psychology,* 1975, **20,** 391–399.

Hinrichs, J. V., & Grunke, M. E. Control processes in short-term memory: Use of retention interval information. *Journal of Experimental Psychology: Human Learning and Memory,* 1975, **104,** 229–237.

Hockey, R. Rate of presentation in running memory and direct manipulation of input-processing strategies. *Quarterly Journal of Experimental Psychology,* 1973, **25,** 104–111.

Imhoff, D. L., Horton, D. L., Weldon, L. J., & Phillips, R. V. Rehearsal and processing capacity as factors in memory. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, **3,** 551–559.

Izawa, C. Vocalized and silent tests in paired-associate learning. *American Journal of Psychology,* 1976, **89,** 681–693.

Jacoby, L. L. Encoding processes, rehearsal, and recall requirements. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 302–310.

Jacoby, L. L., & Bartz, W. H. Rehearsal and transfer to LTM. *Journal of Verbal Learning and Verbal Behavior,* 1972, **11,** 561–565.

Jahnke, J. C., & Davidson, W. R. The effects of three temporal variables on short-term memory for paired associates. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6,** 946–949.

Jeffery, R. W. The influence of symbolic and motor rehearsal in observational learning. *Journal of Research in Personality,* 1976, **10,** 116–127.

Johansson, B. S. Encoding strategy and retroactive interference in short-term memory. *Scandanavian Journal of Psychology,* 1975, **16,** 319–322.

Johnson, M. K., Raye, C. L., Wang, A. Y., & Taylor, T. H. Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning and Memory,* 1979, **5,** 229–240.

Johnson, R. E. Rehearsal effectiveness as a function of similarity of rehearsal activities to a paired-associate task. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7,** 439–445.

Johnson, R. E. Differential availability of associative components in rehearsal. *Journal of Experimental Psychology,* 1970, **85,** 356–360.

Johnson, R. E. Rehearsal of associations stored in long-term memory. *Memory and Cognition,* 1975, **3,** 536–540.

Johnston, W. E., Greenberg, S. N., Fisher, R. P., & Martin, D. W. Divided attention: A vehicle for monitoring memory processes. *Journal of Experimental Psychology,* 1970, **83,** 164–171.

Jongeward, R. H., Jr., Woodward, A. E., Jr., & Bjork, R. A. The relative roles of input and output mechanisms in directed forgetting. *Memory and Cognition,* 1975, **3,** 51–57.

Kahneman, D., & Wright, P. Changes of pupil size and rehearsal strategies in a short-term memory task. *Quarterly Journal of Experimental Psychology,* 1971, **23,** 187–196.

Keele, S. W., & Boies, S. J. Processing demands of sequential information. *Memory and Cognition,* 1973, **1,** 85–90.

Kestner, J., & Walter, D. A. Awareness and the effect of rote rehearsal on free recall. *American Journal of Psychology,* 1977, **90,** 29–37.

Kroll, N. E. A., & Kellicutt, M. H. Short-term recall as a function of covert rehearsal and of intervening task. *Journal of Verbal Learning and Verbal Behavior,* 1972, **11,** 196–204.

Kroll, N. E. A., Kellicutt, M. H., & Parks, T. E. Rehearsal of visual and auditory stimuli while shadowing. *Journal of Experimental Psychology: Human Learning and Memory,* 1975, **1,** 215–222.

Lachman, R., & Mistler, J. L. Rehearsal, test trials, and component processes in free recall. *Journal of Experimental Psychology,* 1970, **85,** 374–382.

Leicht, K. L. Differential rehearsal and primacy effects. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7,** 1115–1117.

Levin, J. R., & Ghatala, E. S. More about rehearsal strategy effects in children's discrimination learning. *Journal of Experimental Psychology: Human Learning and Memory,* 1976, **2,** 381–390.

Locke, J. L., & Ginsberg, M. Electromyography and lipreading in the detection of verbal rehearsal. *Bulletin of the Psychonomic Society,* 1975, **5,** 246–248.

Lockhead, G. R. A re-evaluation of one-trial associative learning. *American Journal of Psychology,* 1961, **74,** 590–595.

Longstreth, L. E. Relationship between response learning and recall of feedback in tests of the law of effect. *Journal of Experimental Psychology,* 1971, **90,** 149–151.

McBane, B. M. Rehearsal capacity and dimensional independence in retardates. *Journal of Experimental Child Psychology,* 1976, **22,** 216–228.

McDaniel, M. A., & Masson, M. E. Long-term retention: When incidental processing fails. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, **3,** 270–281.

McFarland, C. E., Jr., & Kellas, G. Mode of input effects on subject-controlled processes. *Journal of Experimental Psychology,* 1974, **103,** 343–350.

MacLeod, C. M. Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory,* 1975, **104,** 271–279.

McVaugh, W. H. Effects of differences in stimulus retention on concept identification. *Journal of Experimental Psychology,* 1973, **99,** 369–374.

Marshall, P. H., & Werder, P. R. The effects of the elimination of rehearsal on primacy and recency. *Journal of Verbal Learning and Verbal Behavior,* 1972, **11,** 649–653.

Melton, A. W. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 596–606.

Meunier, G. F., Kestner, J., Meunier, J. A., & Ritz, D. Overt rehearsal and long-term retention. *Journal of Experimental Psychology,* 1974, **102,** 913–914.

Meunier, G. F., Meunier, J. A., & Ritz, D. Postlist rehearsal and the recency effect of serial recall. *Journal of Psychology,* 1971, **79,** 299–303.

Meunier, G. F., Ritz, D., & Meunier, J. A. Rehearsal of individual items in short-term memory. *Journal of Experimental Psychology,* 1972, **95,** 465–467.

Modigliani, V. Effects on a later recall by delaying initial recall. *Journal of Experimental Psychology: Human Learning and Memory,* 1976, **2,** 609–622.

Montague, W. E., Hillix, W. A., Kiess, H. O., & Harris, R. Variation in reports of covert rehearsal and in STM produced by differential payoff. *Journal of Experimental Psychology,* 1970, **83,** 249–254.

Monty, R. A., Karsh, R., & Taub, H. A. Pacing of rehearsal in sequential short-term memory. *Journal of Experimental Psychology,* 1967, **74,** 300–302.

Morris, W. (Ed.) *The American heritage dictionary of the English language.* Boston: Houghton, 1975.

Moss, S. M., & Sanders, A. F. The effect of an additional task on recall accuracy in serial short-term memory: Preemption of rehearsal or acquisition? *Acta Psychologica,* 1973, **37,** 307–313.

Moss, S. M., & Sharac, J. Accuracy and latency in short-term memory: Evidence for a dual retrieval process. *Journal of Experimental Psychology,* 1970, **84,** 40–46.

Murdock, B. B., Jr. Effects of a subsidiary task on short-term memory. *British Journal of Psychology,* 1965, **56,** 413–419.

Murdock, B., & Metcalfe, J. Controlled rehearsal in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior,* 1978, **17,** 309–324.

Murray, D. J. Intralist interference and rehearsal time in short-term memory. *Canadian Journal of Psychology,* 1966, **20,** 413–426.

Murray, D. J. Overt versus covert rehearsal in short-term memory. *Psychonomic Science,* 1967, **7,** 363–364.

Neimark, E., Greenhouse, P., Law, S., & Weinheimer, S. The effect of rehearsal-preventing task upon retention of CVC syllables. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4,** 280–285.

Neisser, U. *Cognitive psychology*. New York: Appleton, 1967.

Nodine, C. F. Temporal variables in paired-associate learning: The law of contiguity revisited. *Psychological Review*, 1969, **76**, 351-362.

Nodine, C. F., Nodine, B. F., & Thomas, R. C. Temporal variables in paired-associates learning: The roles of repetition and number tracking during stimulus intervals. *Journal of Experimental Psychology*, 1967, **73**, 439-445.

Olton, R. M., & Johnson, D. M. Mechanisms of incubation in creative problem solving. *American Journal of Psychology*, 1976, **89**, 617-630.

Palmer, S. E., & Ornstein, P. A. Role of rehearsal strategy in serial probed recall. *Journal of Experimental Psychology*, 1971, **88**, 60-66.

Penney, C. G. Modality effects in short-term verbal memory. *Psychological Bulletin*, 1975, **82**, 68-84.

Peterson, L. R. Concurrent verbal activity. *Psychological Review*, 1969, **76**, 376-386.

Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, **58**, 193-198.

Peterson, M. J., Peterson, L. A., & Ward-Hull, C. Rehearsing and retaining mental matrices. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 371-381.

Peterson, M. J., Thomas, J. E., & Johnson, H. Imagery, rehearsal, and the compatibility of input-output tasks. *Memory and Cognition*, 1977, **5**, 415-422.

Petrusic, W. M., & Jamieson, D. G. Differential interpolation effects in free recall. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, **4**, 101-109.

Pollack, I. Interference, rehearsal, and short-term retention of digits. *Canadian Journal of Psychology*, 1963, **17**, 380-392.

Pollatsek, A., & Bettencourt, H. O. The spaced-practice effect in the distractor paradigm is related to proactive interference but not to short-term store. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, **2**, 128-141.

Posner, M. E., & Rossman, E. Effect of size and location of informational transforms upon short-term retention. *Journal of Experimental Psychology*, 1965, **70**, 496-505.

Postman, L., & Kruesi, E. The influence of orienting tasks on the encoding and recall of words. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 353-369.

Postman, L., & Phillips, L. W. Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology*, 1965, **17**, 132-138.

Postman, L., & Warren, L. Test of the total-time hypothesis in free-recall learning. *Journal of Experimental Psychology*, 1972, **96**, 176-183.

Poulton, E. C. Continuous intense noise masks auditory feedback and inner speech. *Psychological Bulletin*, 1977, **84**, 977-1001.

Reed, H. Studies of the interference process in short-term memory. *Journal of Experimental Psychology*, 1970, **84**, 452-457.

Reitman, J. S. Mechanisms of forgetting in short-term memory. *Cognitive Psychology*, 1971, **2**, 185-195.

Reitman, J. S. Without surreptitious rehearsal, information in short-term memory decays. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 365-377.

Reitman, W., Malin, J. T., Bjork, R. A., & Higman, B. Strategy control and directed forgetting. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 140-149.

Reynolds, J. H. Confirmation, contiguity, and response practice in paired-associate learning. *Journal of Experimental Psychology*, 1967, **73**, 394-400.

Roediger, H. L., III, & Crowder, R. G. Instructed forgetting: Rehearsal control or retrieval inhibition (repression)? *Cognitive Psychology*, 1972, **3**, 244-254.

Roenker, D. L. Role of rehearsal in long-term retention. *Journal of Experimental Psychology*, 1974, **103**, 368-371.

Rowe, E. J. The effect of imagery on rehearsal strategies in verbal-discrimination learning. *American Journal of Psychology,* 1975, **88,** 431-442.

Rubin-Rabson, G. Mental and keyboard overlearning in memorizing piano music. *Journal of Musicology,* 1941, **3,** 33-40.

Rundus, D. Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology,* 1971, **89,** 63-77.

Rundus, D. Output order and rehearsal in multi-trial free recall. *Journal of Verbal Learning and Verbal Behavior,* 1974, **13,** 656-663.

Rundus, D. Maintenance rehearsal and single-level processing. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 665-681.

Rundus, D., & Atkinson, R. C. Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 99-105.

Rundus, D., Loftus, G. R., & Atkinson, R. C. Immediate free recall and three-week delayed recognition. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 684-688.

Ryan, J. Temporal grouping, rehearsal and short-term memory. *Quarterly Journal of Experimental Psychology,* 1969, **21,** 148-155.

Sackett, R. S. The relationship between amount of symbolic rehearsal and retention of a maze habit. *Journal of General Psychology,* 1935, **12-13,** 113-128.

Sanders, A. F. Rehearsal and recall in immediate memory. *Ergonomics,* 1961, **4,** 23-34.

Santa, J. L., & Ranken, H. B. Effects of verbal coding on recognition memory. *Journal of Experimental Psychology,* 1972, **93,** 268-278.

Sassenrath, J. M. Theory and results on feedback and retention. *Journal of Educational Psychology,* 1975, **67,** 894-899.

Schendel, J. D. Analysis of the rehearsal processes in the release from proactive interference. *Journal of Experimental Psychology: Human Learning and Memory,* 1976, **2,** 76-82.

Seamon, J. G., & Wright, C. E. Generative processes in character classification: Evidence for a probe encoding set. *Memory and Cognition,* 1976, **4,** 96-102.

Shaffer, W. O., & Shiffrin, R. M. Rehearsal and storage of visual information. *Journal of Experimental Psychology,* 1972, **92,** 292-296.

Shaughnessy, J. J., & Underwood, B. J. The retention of frequency information for categorized lists. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 99-107.

Shebilske, W., & Epstein, W. Effect of forget instructions with and without the conditions for selective search. *Memory and Cognition,* 1973, **1,** 261-267.

Shebilske, W., Wilder, L., & Epstein, W. Forget instructions: Effect of selective rehearsal and categorical distinctiveness. *Journal of Experimental Psychology,* 1971, **89,** 372-378.

Shiffrin, R. M. Information persistence in short-term memory. *Journal of Experimental Psychology,* 1973, **100,** 39-49.

Smyth, M. M. The role of mental practice in skill acquisition. *Journal of Motor Behavior,* 1975, **7,** 199-206.

Sperling, G. The information available in brief visual presentations. *Psychological Monographs,* 1960, **74,** No. 11 (Whole No. 498).

Sperling, G., & Speelman, R. G. Acoustic similarity and auditory short-term memory: Experiments and a model. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970.

Stanners, R. F., Meunier, G. F., & Headley, D. B. Reaction time as an index of rehearsal in short-term memory. *Journal of Experimental Psychology,* 1969, **82,** 566-570.

Starr, B. J., Harris, G. J., & Aronoff, R. Variable length consonant stimuli in a "Peterson & Peterson" short-term memory paradigm. *Psychonomic Science,* 1970, **20,** 243-244.

Stelmach, G. E., & Bassin, S. L. The role of overt motor rehearsal in kinesthetic recall. *Acta Psychologica,* 1971, **35,** 56-63.

Sternberg, S. Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist,* 1969, **57,** 421-457.

Sternberg, S. Memory-scanning: Mental processes revealed by reaction-time experiments. In J. S. Antrobus (Ed.), *Cognition and affect.* Boston: Little, Brown, 1970.

Tarpy, R. M., Glucksberg, S., & Lytle, L. D. Motivation and short-term retention: Evidence for covert rehearsal. *American Journal of Psychology,* 1969, **82,** 111-116.

Ternes, W., & Yuille, J. C. Words and pictures in an STM task. *Journal of Experimental Psychology,* 1972, **96,** 78-86.

Thompson, J. T., & Clayton, K. N. Presentation modality, rehearsal-prevention conditions, and auditory confusions in tests of short-term memory. *Memory and Cognition,* 1974, **2,** 426-430.

Timmins, W. K. Repetition of intentionally forgotten items. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 168-173.

Turvey, M. T. The effects of rehearsing analyzed information upon the retrieval of unanalyzed information. *Psychonomic Science,* 1966, **6,** 365-366.

Tzeng, O. J. L. A precedence effect in the processing of verbal information. *American Journal of Psychology,* 1976, **89,** 577-599.

Tzeng, O. J. L., & Hung, D. L. Intralist organization and subsequent free recalls. *Journal of Experimental Psychology,* 1973, **98,** 119-124.

Ulich, E. Some experiments on the function of mental training in the acquisition of motor skills. *Ergonomics,* 1967, **10,** 411-419.

Underwood, B. J. A breakdown of the total-time law in free-recall learning. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 573-580.

Underwood, B. J. *Temporal codes for memories: Issues and problems.* Hillsdale, N.J.: Erlbaum, 1977.

Underwood, B. J., & Keppel, G. An evaluation of two problems of method in the study of retention. *American Journal of Psychology,* 1962, **75,** 1-17.

Underwood, B. J., & Postman, L. Extraexperimental sources of interference in forgetting. *Psychological Review,* 1960, **67,** 73-95.

Wagner, A. R., Rudy, J. W., & Whitlow, J. W. Rehearsal in animal conditioning. *Journal of Experimental Psychology,* 1973, **97,** 407-426.

Weiner, B., & Reed, H. Effects of the instructional sets to remember and to forget on short-term retention: Studies of rehearsal control and retrieval inhibition (repression). *Journal of Experimental Psychology,* 1969, **79,** 226-232.

Weist, R. M. The role of rehearsal: Recopy or reconstruct. *Journal of Verbal Learning and Verbal Behavior,* 1972, **11,** 440-450.

Wetzel, C. D. Effect of orienting tasks and cue timing on the free recall of remember- and forget-cued words. *Journal of Experimental Psychology: Human Learning and Memory,* 1975, **1,** 556-566.

Wetzel, C. D., & Hunt, R. E. Cue delay and the role of rehearsal in directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, **3,** 233-245.

Whitten, W. B., II, & Bjork, R. A. Learning from tests: Effects of spacing. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 465-478.

Wickelgren, W. A. Rehearsal grouping and hierarchical organization of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology,* 1967, **19,** 97-102.

Woodward, A. E., Jr., Bjork, R. A., & Jongeward, R. H., Jr. Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 608-617.

Woodward, A. E., Jr., Park, D. C., & Seebohm, K. Directed forgetting as a function of explicit within-list cuing and implicit postlist cuing. *Journal of Experimental Psychology,* 1974, **102,** 1001-1006.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology* (rev. ed.), New York: Holt, 1954.

INDIVIDUAL DIFFERENCES IN FREE RECALL: WHEN SOME PEOPLE REMEMBER BETTER THAN OTHERS[1]

Marcia Ozier

DALHOUSIE UNIVERSITY
HALIFAX, NOVA SCOTIA, CANADA

---

[1] The preparation of this chapter and the experiments reported herein have been supported in part by Grant APA-214 from the National Research Council of Canada. The data have been gathered over a 12-year period by a coterie of admirable undergraduate research assistants who were (in chronological order): Beatrice Smith, Sandra Brown, Susan Hughes, Debi Mitchell, Martha Castell, Elizabeth Thompson, Ann Marie Long, Ruth Evans, Jamie Stewart, Adele Blennerhassett, Margaret Matthews, Dwain Campbell; two of these, Ellen Wood and Bruce Clarke deserve special mention for the imagination they applied to the extensive analysis of the data they collected. Several of the experiments were conducted as honors, qualifying, or master's theses. Here the contributions of Orville Endicott, Josie Kickam, Bruce Anderson, and John Sperry are gratefully acknowledged; they are the "we" in this chapter. The subjects in these experiments have been first-year undergraduates not only at the University of Toronto and Dalhousie, but also at St. Mary's, Mount St. Vincent, and Acadia Universities. The cooperation of all of these institutions was helpful; special thanks are due to Sister Mary Albertus of Mount St. Vincent. Finally, the author acknowledges the special contributions of Nancy Beattie who prepared the final manuscript with great care and good humor, Bruce Earhard who provided valuable and sensible advice and unlimited assistance throughout the preparation of the manuscript, and Endel Tulving who sparked her curiosity about IDs in memory many years ago.

THE PSYCHOLOGY OF LEARNING
AND MOTIVATION, VOL. 14

ISBN 0-12-543314-X

I. The Problem

The exercise of observing the individuals in a group of college students committing a list of unrelated words to memory under conditions of multitrial free recall is an extremely instructive one. Although the performance of these subjects is usually reported for the group as a whole, there are as many styles of memorizing as there are memorizers. Some students are calm; some are tense. Some write seemingly effortlessly; some put their whole beings into the output of a word. Some write fluidly, one word following another without hesitation; some write in fits and starts, pausing or not between words. Some write more and more words as the trials continue; some write fewer and fewer after an initial spurt. Some write the same words over and over; some write different words each trial. And these, of course, are only the most superficial differences. One must assume that the subjects differ, also, in a variety of mental processes brought to bear on the task set before them, and in the strategies (the cognitive style) they use—in order to master the material.

The importance of understanding individual differences (IDs) in memory and the contribution that such understanding might make to our general theories of cognitive function has not gone unnoticed by experimental psychologists (e.g., Gagne, 1967; Underwood, 1975). Earlier, Jenkins (1961) had spoken of the [regrettable] neglect of IDs in most laboratories; in a more positive tone, Postman (1961) had pointed out the potential for using individual differences to make discoveries about the learning process itself, a position held also by Sakoda (1956) who stressed the importance of IDs for "increased insight into the effects of the experimental variables themselves" (p. 189). In fact, none of the 21 distinguished contributors to the Gagne volume denied the tremendous potential that research focused upon IDs had for advancement of our knowledge about how the human information processing system operates (although Cofer did state that he remained to be convinced). Yet, despite the importance and the potential and the challenge recognized by all, sustained, systematic, and imaginative attacks upon this problem have been very few and far between.

There have been some noble attempts however. One beginning was offered by Postman and his colleagues (e.g., Plenderleith & Postman, 1957) who tested the hypothesis that retention was a function of the strength (defined as conventionality) of a subject's differential responses to the to-be-remembered materials. Plenderleith and Postman concluded with the optimistic statement that "individual differences in selective retention can be related to the specific verbal dispositions with which the learner enters the experimental situation" (p. 248). In a more modern vein, Dean and Ley (1977) have asked whether the number of associations subjects give to verbal units can predict their free-recall performance. Their results have encouraged them to consider seriously associative ability as an individual difference variable (Ley & Dean, 1976; Locascio & Ley, 1972). Whereas Postman considered the conventionality of the associative response and Ley considered the number of associative responses, Earhard (1967a, 1970a, 1974) and Earhard and Endicott (1969) attributed free-recall IDs to the ability to form and maintain inter-item associations. This hypothesis had seemed to Jenkins "the most obvious candidate for an important variable . . . the postulated ability to form new associative bonds or habits" (1967, p. 48). With these few exceptions, however, in the years since 1967, the issue of why some people memorize better than others has remained of much general interest but of not much particular import.

The same seems somewhat less the case for the pursuit of individual differences in shorter term memory processes. In recent publications, Jensen (1971) has asked whether there are visual and auditory types of memories; the answer is no. Chiang and Atkinson (1976) have asked whether IDs in memory search and visual search tasks are related; the answer seems to be yes if one takes the component processes and the sexes of the subjects separately (but see Gagnon, Cavanagh, & Laurencelle, 1978, for another opinion). And Lyon (1977) has explored the questions of whether IDs in immediate serial recall are attributable to mnemonic (grouping, chunking, rehearsal); the answer is no. A major systematic and important contribution to this area has been made by Hunt and his associates (Hunt, Frost, & Lunneborg, 1973; Hunt, Lunneborg, & Lewis, 1975; McLeod, Hunt, & Mathews, 1978) in which the issue of what it means to be high verbal has been addressed, the conclusion being that "although a verbal intelligence test is directly a measure of what people know, it is indirectly a way of identifying people who can code and manipulate verbal stimuli rapidly in situations in which knowledge *per se* is not a major factor" (p. 223). Moreover, Hunt (1978) has addressed the theory of individual differences in cognition pointing out three sources of these: (*a*) knowledge, (*b*) mechanical capability, and (*c*) general information processing techniques. We will be using these sources as a framework for the final discussion of the data presented later in this chapter.

Finally, there are some individual differences which have interested psychologists a lot. Eysenck (1977) collating current contributions to the individual-differences literature has included chapters on arousal, introversion–extroversion, anxiety and neuroticism, aging, and intelligence as related to human-memory function. The conclusions to be drawn about these dimensions of human nature as they influence cognitive processes are not clear. What is clear and what will be assumed below is that in the absence of wide variations in these variables, there are still IDs in multitrial free recall, differences that must be sought somewhere in the complexities of the mental processing of stimuli arriving from our external and internal worlds.

II. The Subjects

This article reports the results of numerous studies each one designed to test a hypothesis about why some people do better than others under conditions of multitrial free recall. That the range of performance within the narrow span of humanity represented by college freshmen is wide can be demonstrated very easily. We now have data for thousands of students who performed for us exactly the same task. Twenty-two unrelated English words (there are several sets of these) are presented one at a time at the rate of 1 sec per word, and after each series of presentations the subjects' task is to recall, in writing, as many words as they can remember in any order they like. The same words are presented (but their order is changed) and recalled during 16 trials; subjects write their recall either on paper pullers which permit only one word to be seen at any moment or using a cardboard mask with which they cover the items they have already written.

The distribution of the last 157 subjects' performance scores is shown in the histogram of Fig. 1 for the average performance over 16 trials, and for Trials 1 and 16. Here one can see that, on the average, subjects recall about eight words on Trial 1, 19 words on Trial 16, and 15 words per trial, over all. Although these distributions are approximately normal and symmetrical, in all cases the variability in performance is striking, as one sees, for example, 22 subjects who averaged fewer than 13 items per trial and 29 subjects who averaged 18 or better.

How is one to account for such a range of performance scores? The first and most natural answer and one which always is mentioned has to do with "intelligence." The reason must be that the good memorizers are smarter than the poorer ones—smarter in some general sense. That this is not the case has been argued forcefully by Woodrow (1946), and our own data support his statement that "The ability to learn cannot be identified with the ability known as intelligence" (p. 148). We have administered Raven's Progressive Matrices, often considered a test of the most general aspect of intelligence (notwithstanding Hunt, 1974), and the Shipley Institute of Living Scale (Shipley, 1940) a quick, self-administered, paper

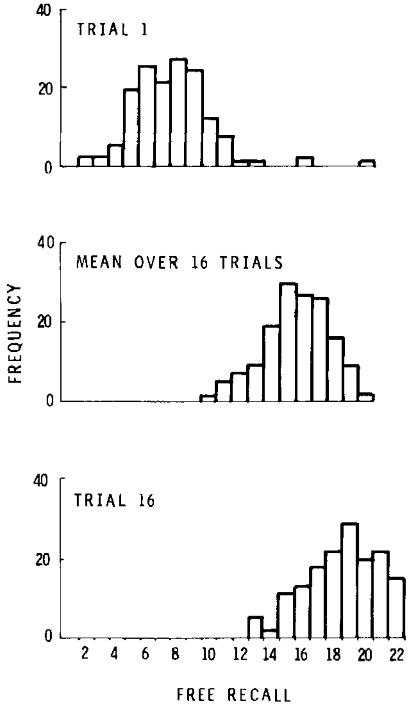
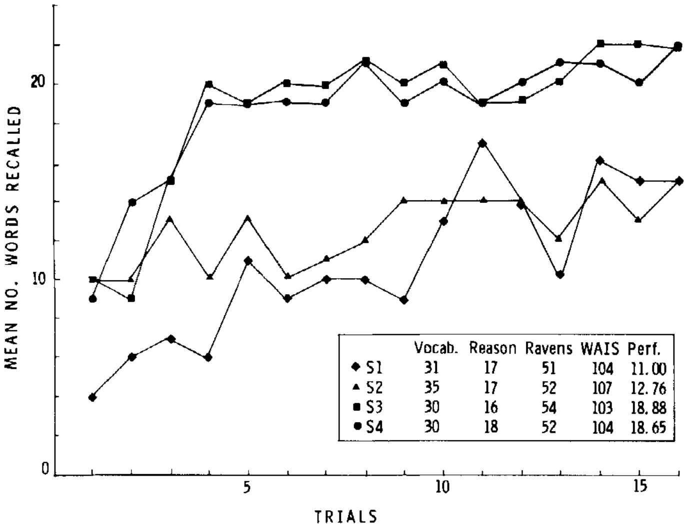


Fig. 1. The frequency with which various levels of free-recall performance was obtained by subjects during Trial 1, Trial 16, and the average over 16 trials.

and pencil test designed to aid in the detection of mild intellectual impairment in persons of normal original intelligence. The Shipley test consists of a vocabulary test and a test of abstract reasoning. We have also measured the digit span of large groups of subjects. Groups of 24 or more subjects selected randomly from the worst performers and the best performers had similar digit span, similar average Raven's scores (53.0 and 54.3), similar Shipley vocabulary scores (29.8 vs 31.0), similar Shipley abstraction scores (17.2 vs 17.9), and the WAIS equivalents of their Shipley totals were 104 and 105, respectively.

Yet the curves these subjects produce in attempting to memorize a list of unrelated items are grossly different, as can be seen in Fig. 2, where we have presented the acquisition curves of four subjects of average scores on the Raven and both Shipley scales and the WAIS score. Our results have indicated that, as Woodrow knew 30 years ago, the simplest hypothesis to explain the behavior gap between subjects 1 and 2 vs 3 and 4 is simply not tenable. And so we will assume as this discussion develops that what allows the good memorizers to remember better is not



	Vocab.	Reason	Ravens	WAIS	Perf.
◆ S1	31	17	51	104	11.00
▲ S2	35	17	52	107	12.76
■ S3	30	16	54	103	18.88
● S4	30	18	52	104	18.65

Fig. 2. Mean number of words correctly recalled for four subjects with their scores on the Vocabulary Test, Reasoning Test, Raven's Progressive Matrices, WAIS, and average performance over 16 trials.

something that they *are,* like "are smarter" for example, but something that they *do.*

III. Organization and Memory

"The great thing, then," wrote William James (1892, p. 144) "is to *make our nervous system our ally instead of our enemy.*" There is probably wide agreement that one of the aspects of the nervous system to be classed as the "enemy" is the memory system. It is doubtful that even the most talented mnemonist would be prepared to consider the memory system as an "ally." But the psychological literature since 1892, and particularly since 1953, does tell us much about the apparent limits of the "enemy" to remembering, and we now recognize a number of the activities in which we engage in order to overcome those limits. This article will concentrate upon one of those activities: the finding that, as trials proceed, the order of recall becomes more and more fixed (Tulving, 1962) reflecting the imposition of an organization or structure (Bower, Lesgold, & Tieman, 1969; Buschke, 1976; Friendly, 1977) upon the material being committed to memory. To quote James again, "Briefly, then, of two men with the same outward tenacity, the one who THINKS over his experiences most, and weaves them into systematic relations with each other, will be the one with the best memory" (p. 294).

There are differences of opinion as to the best measure of the subjective sequential structure reflected in the order of output of free recall. These are presented in detail by Sternberg and Tulving (1977; but see also Friendly, 1977) who find that the best measure is bidirectional pair frequency (PF), one of a class of measures that reflect "constancies in output order, under conditions wherein such constancies are not required of subjects' performance, and which cannot be readily attributed to processes other than utilization of elementary units into higher order ones. The degree of output consistency over trials can thus be used as an index to which a particular organization has occurred and is maintained from one trial to the next" (p. 540). PF and recall performance are highly correlated (.78), making this a relatively reliable measure of organization of recall. An earlier measure of sequential or subjective organization was SO2 (Tulving, 1962), a statistic not as reliable in predicting recall performance (.60 in our data) but highly correlated with PF (.80 for our subjects). All of the experiments reported below used subjects' PF or SO2 or both as IDs in subjective organization.

While the idea that memorization proceeds by organization imposed upon unordered material which increases with recall makes good intuitive sense (but see Landauer, 1975, for an interesting alternative), the relevance of organization that is reflected in SO2 has been questioned by some (e.g., Gorfein, Blair, & Rowland, 1968; Shapiro & Bell, 1970). The argument has been that the sequential organization observed seems relatively minimal compared to the organization that could have taken place and the recall that does take place. Also, subjects who organize poorly do in fact acquire some of the list and free recall improves with age with not much increase in subjective organization (Ornstein, Naus, & Liberty, 1975). We would give two kinds of replies to this kind of doubt. First, and as Sternberg and Tulving argue for us, the measures of organization reflect the minimum organization subjectively imposed. There is no claim that the measures capture the whole structural process. And so it is not that surprising that recall proceeds at a different rate than organizational measures increase. Second, we would not wish to claim that subjective organization is the only variable controlling increases in retention with multiple trials. Finally, while it is the case that a few subjects who seem to be able to remember as many as 12 or 14 words without any attention to the serial order in which they are reported, it is also true that the pattern of their behavior is somewhat different. It appears as though an asymptote for retrieval is reached very early by these subjects, and that performance continues at that level for trial after trial, regardless of the fact that there may be 10 or more items left to incorporate into recall in some fashion or other. Notwithstanding the above, for most subjects, IDs in organization certainly seem to be related to the ultimate success with which items are committed to memory. In fact, it is argued strongly by some (e.g., Buschke, 1976, 1977) that, for optimum recall, subjective organization *must* take place (see also Bellezza, Richards, & Geiselman, 1976). It is important to note, moreover, that the effects of IDs in organization develop over trials (Earhard, 1974).

The experiments reported below tested several hypotheses concerning IDs in subjective organization. The underlying assumption has been that subjective organization reflects the sequential structure imposed upon the materials to be committed to memory by the average subject and so reflects the normal functioning of the mental apparatus when faced with multitrial free recall. Our approach to this research has been (to coin a phrase) "relative naiveté," otherwise known as the extreme-groups method. Subjects who participated in our multitrial free-recall procedures were classified as high or low organizers (LS or HS) on the basis of whether their organization scores (PF and/or SO2) fell below or above the mean score of the complete distribution of scores collected.

Figure 3 shows the distributions of PF and SO2 scores for the subjects whose recall we saw in Fig. 1. The mean SO2 was .32 (s = .063), the mean PF was 2.34 (s = 1.517). The correlation between SO2 and recall performance over 16 trials was +.59, and for PF was +.75; the correlation between SO2 and PF was +.91.

Following the classification of subjects as LS or HS, equal numbers of these have participated in studies designed to reveal interactions between independent variables and our ID variable (see Hultsch, 1971; Jablonski, 1974; Jacoby, Bartz, & Evans, 1978). (All of our subjects participated then in *two* separate sessions. The first to establish SO2/PF; the second to test whatever current hypothesis was being examined.) The remote purpose of these studies was always to explain why some people are better subjective organizers than others—in order to explain why some people remember better than others.

IV. The Null Hypothesis Accepted

This first section summarizes experiments in which we have found no differences between LS and HS. There have been only three occasions on which this result has been obtained: in short-term memory for consonant syllables (Earhard, 1970b; à la Peterson & Peterson, 1959), in latency of retrieval of instances from semantic memory, and the reflected spread of

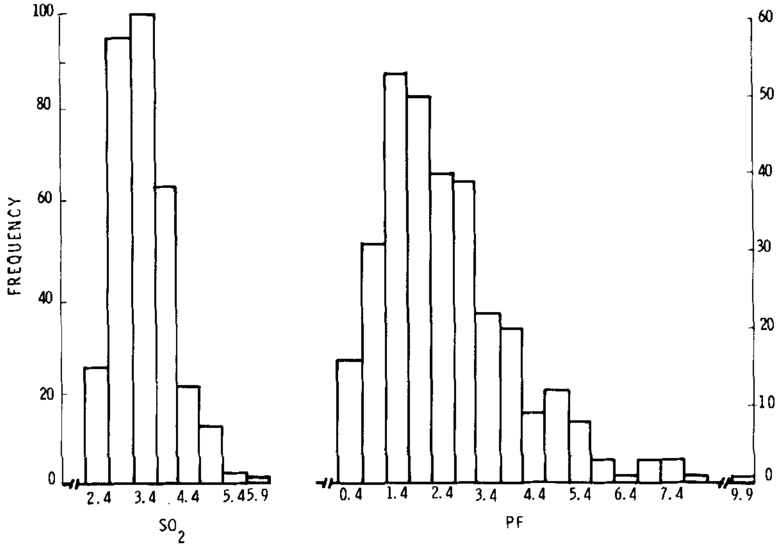


Fig. 3. Frequency distribution of the number of subjects obtaining SO2 and PF scores of various values.

activation, given the category name and initial letter specifications (Freedman & Loftus, 1971; Loftus, 1973; Loftus & Loftus, 1974), and in digit span after Primary (Type I) Rehearsal (Glenberg & Adams, 1978).

The first of these studies (Earhard, 1970b) examined the hypothesis that differences between LS and HS originated in IDs in the ability to retain each item in the set independent of the associative and serial structure imposed. To examine this matter LS and HS were presented a series of 48 Witmer consonant syllables (Hilgard, 1951), each of which was followed by a three-digit number from which subjects counted backward by threes until 0, 3, 6, 9, 12, or 18 sec had passed. The subject's task at that point was to report the most recent consonant syllable. The results are presented in Table I where it is obvious that short-term retention of this type did not vary as a function of subjective organization classification, $F(1,28) < 1$. Whatever it is that the high organizer does better to help him/her during multitrial free recall, that activity does not seem to operate during classic retention over short intervals.

While the data reported above come from an experiment in which the retention interval was filled with maximum-effort activity, we also have data from a replication of a design used by Glenberg and Adams (1978) in which the short-term retention interval was filled with minimum-effort rehearsal. In brief, the Glenberg and Adams procedure requires subjects

TABLE I

PROBABILITY OF CORRECT RECALL AS A FUNCTION OF THE RETENTION INTERVAL FOR LS AND HS

	Retention interval in seconds				
	0	3	6	9	12
LS	.89	.55	.37	.45	.40
HS	.92	.62	.36	.39	.26

to retain four-digit numbers over an interval during which a pair of words is rehearsed overtly 1, 5, or 10 times. The primary data collected are from a surprise posttest in which recognition of the rehearsed words is required, and these will be described in detail later on. The data of interest here are those representing the retention of the digits, over the rehearsal interval, which was not different for LS and HS, the mean errors being 17.4 and 15.8, a difference which was not statistically significant, $F(1,28) = 1.45$, $p > .05$. It seems as though the short-term retention tasks we have chosen have failed to discriminate between high and low subjective organizers; this aspect of the memory system is probably not the locus of the primary difference(s) between them.

The third instance of there being no difference between LS and HS occurred in an experiment designed to examine the nature of the subjective lexicon (Miller, 1970), that cognitive subsystem which corresponds to our knowledge about words: their names and their meanings. The design of our experiment was taken from the work of Loftus on the spread of activation in semantic memory (Loftus, 1973; Loftus & Cole, 1974) in which subjects are asked to respond to a category-letter pair with an instance that belongs to the category and begins with the letter, and then, immediately or after one or two other trials have occurred, the original category is repeated but with a different letter, eliciting a different instance. The reduced latency of response to the second occurrence of the category name has been taken as evidence for the spread of activation within categories (Collins & Loftus, 1975) although Hopf-Weichel (1977) had proposed a very interesting alternative interpretation in terms of the reorganization of items to and from active states. Notwithstanding differences in interpretation (see also Keller & Kellas, 1978), the Loftus paradigm provided a neat method for testing whether LS and HS differ in the speed with which they could access items in semantic memory.

Our experiment included two other comparisons as well as the category same-letter different (fruit-p, fruit-a) repetition. In order to examine the spread of activation in the lexical dictionary of the *names* of words, we included occasions when the category names were different, but the initial letters were the same (insect-a, fruit-a); and, as a control for both of these, we included a condition, occasions when the same category and the same letter were repeated (fruit-a, fruit-a). These materials had been used in our laboratory previously (Ozier, 1979) to examine developmental differences in spread of activation of semantic and phonemic aspects of words. The subjects' task was to respond as quickly as possible with an instance that belonged to the category and began with the letter shown, and the reaction time measures were taken by a voice key and recorded by computer.

The results of this study were as follows. LS and HS subjects did not differ in the initial presentations (212 vs 212 msec), not in the control condition (category same/letter same: 179 vs 167 msec), not in the category same/letter different condition (191 vs 194 msec), and not in the category different/letter same condition (220 vs 225 msec). Neither were there any interactions with lag. In view of the clear outcome of this experiment, one might, with some degree of confidence, accept the null hypothesis that LS and HS subjects perform no differently in the speed or the accuracy with which they access information in the mental dictionary.

V. The Original Experiments

The early experiments in this series did show consistent, reliable, and intriguing differences between LS and HS (Earhard, 1967a, 1970b, 1974; Earhard & Endicott, 1969). In the original study, the memorization of a 16-word list over the course of 16 trials was examined for LS and HS under free-recall and under serial-recall instructions. The free recall allowed, of course, for recall in any order the subject chose; the serial recall insisted that the subjects recall the words in the order in which they were presented. The underlying rationale of the experiment was that HS might do better than LS in free recall; but, under conditions of serial recall, where *subjective* organization was not permitted for the HS, and where LS would be forced to recall the items in the same order from trial to trial (as part of the task requirement), it was expected that the differences in performance between these two groups of learners would disappear.

There was a third variable in this experiment besides recall instructions and ID in subjective organization, and that was the order in which the items were presented. There were two input orders; some subjects saw a preferred order (PL), and some a nonpreferred order (NPL). The PL presented the items in an order which followed most closely the order of the recall of those items on Trial 16 developed by the subjects of a previous free-recall experiment. In the NPL, the items were presented in a sequence never found in the recall protocols of previous learners. And so we contrasted, in this experiment, three sources of sequential organization in multitrial free recall: subjects' tendency to organize subjectively the items during learning (i.e., LS vs HS), sequential organization inherent in the task instruction (free vs serial recall), and sequential organization inherent in the order of presentation of the stimulus items (PL and NPL). It was expected that the PL and the serial instructions would allow the differences between LS and HS to disappear.

The results could not have been more unexpected. On every comparison, high organizers did better than low organizers—both under sets of instructions and with both orders of presentation. For HS, the instructions made no difference; for LS the serial instructions resulted in a higher level of performance during the latter half of the trials. For HS, the order of presentation made no difference during free recall, but did during serial recall; for LS, the order of presentation preferred by the previous learners resulted in better memorization. The results of these data seemed pretty clear; HS subjects had a facility for committing unrelated words to memory which operated not only when they were free to order the items according to their own whims, not only when the order in which the items were presented was favorable, but also when the order of the items was fixed, and when the items were sequenced in an order that had been avoided by other learners. The interpretation given these data by Earhard (1967a) was: "It may be that... the advantage of the good organizer is the ability to connect any two stimulus items more rapidly and more permanently than the poor organizer" (p. 507).

To evaluate this original hypothesis, the next step we should take is to look at LS vs HS under conditions of practice in which sequence is not at all an issue, and interitem associations are. The data for this comparison came from Earhard and Endicott (1969) who had LS and HS learn double-function paired-associate lists as part of a more extended series of studies. It was reasoned there that, in PA learning, where the order of the pairs is scrambled in each trial, and there is no intertrial interval, the ability to produce or follow sequences should make a minimal contribution to performance scores, and the mastery of the material must be dependent heavily upon formation of interitem associations between stimuli and responses.

The outcome of this experiment was as follows. The mean number of trials to a criterion of one perfect trial was 51.95 for LS and 35.6 for HS, which reflects a difference late in learning; the mean probability of the correct response during the early trials (1 to 15) was .26 for LS and .36 for HS. This difference in the overall analysis of the first 15 trials was significant, $F(1,36) = 7.47$, $p < .01$, as was the interaction between LS/HS and trials indicating that the gap between LS and HS widened as trials proceeded. For example, the probability of a correct response during the first three trials was .11 for LS and .12 for HS; for trials 13 to 15 the corresponding data were .33 and .55. We considered these results strong support for the "IDs in Association Formation" hypothesis.

The final data in this section come from an experiment in which the recognition memory of LS and HS[2] was compared (Earhard, 1970b). The experiment was designed originally with the complete assurance that recognition memory would level all differences between LS and HS. After all, here was a task which it was strongly held, at that time, did not involve organization (Kintsch, 1968, 1970).

The procedural details were as follows. LS and HS subjects were presented the PL and NPL sequences of items at the rate of 1 (three replications for a total of 35 subjects per condition) or 2 (two replications for a total of 20 subjects per condition) sec per item. After the items had been shown, the subjects were given a sheet of paper upon which were typed 56 words in a column. Among those words, in randomly determined positions, were the 14 target words the subjects had seen, along with three distractors for each word, each distractor beginning with the same letter as its target item, and being of approximately the same frequency of occurrence (Thorndike & Lorge, 1944). The subject's task was to cross out the words he/she had seen on the memory drum. That procedure was repeated 10 times; each time this procedure was repeated for 10 trials; each trial the items were presented in the same order (PL or NPL), but the sequence of target and distractor words on the recognition test was different, and for each subject there was a different order of tests.

The results of each replication of the experiment were entered into an analysis of variance of arc sine transformations of the raw data. The analysis considered Rate of Presentation (1 vs 2 sec), Subjective Organization (LS vs HS), Order of Presentation (PL vs NPL), and Trials (10). The same pattern of results was obtained for each replication. There was a triple interaction between Trials, Order of Presentation, and Subjective Organization, e.g., $F(9,648) = 2.05$, $p < .05$, and the main effect of Rate of Presentation was also significant, e.g., $F(1,36) = 3.98$, $p < .05$. That this was a relatively difficult task is attested to by the results shown in Fig. 4 where the source of the interaction seems apparent in each panel.

---

[2] The numbers of LS and HS in the several experiments reported hereafter have varied from 10 to 24 per condition.

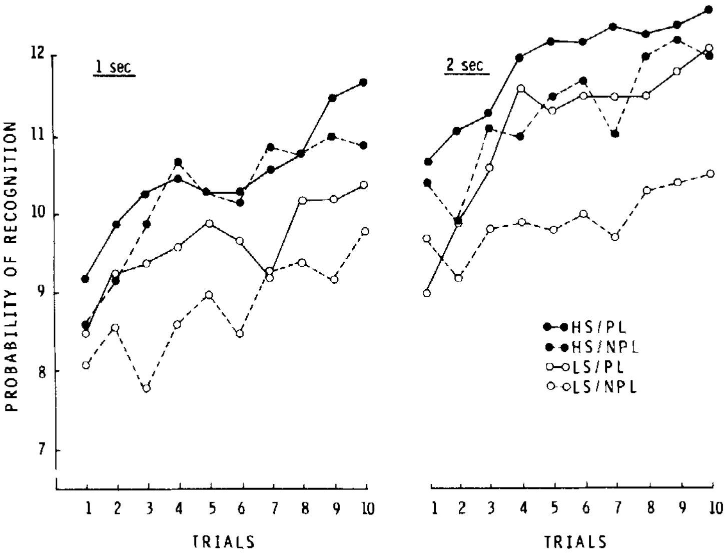


Fig. 4. LS and HS recognition memory as a function of 10 trials for two orders of presentation.

Given the 1- *or* the 2-sec per item rate, HS recognized better than did LS, but LS performance was influenced more by the Order of Presentation variable. LS recognized items better after they were presented in an order that had been preferred for free recall by previous learners; HS recognized almost as well after NPL as after PL sequences.

In retrospect, it should not have been surprising perhaps to find that good memorizers during free recall are better able to recognize also. Tulving (1976), for example, has maintained that "recognition and recall differ only with respect to the exact nature of the retrieval information available to the rememberer" (p. 37). If this is the case, then whatever IDs elevate performance in recognition should elevate performance in recall, and vice versa. One can only speculate about which aspect of recognition and recall might have been unearthed in this study. But the conclusion from these data, that HS advantage during free recall may have nothing to do with creating idiosyncratic sequences per se, is not difficult to draw (see also Earhard, 1974). The advantage of HS in free recall appears likely to be some process or parameter of the memory system which is commonly applicable to free recall, serial recall, paired-associate learning, and recognition memory. [One could entertain seriously, of course, the other alternative, that is, that there are many aspects of HS memories that are superior to those of LS, different ones of these operating in various tasks to produce the consistent differences we have seen (Battig, 1979). There is no final refutation of the argument. It leads, however, to despair; whereas the intuition that, whatever it is that HS do better, is similar in a variety of situations, leads to numerous experiments, some of which are described below.]

The next sections of this chapter describe the pursuit of three hypotheses generated in an effort to reveal which parameter of the memory system might be involved. First, we explored the notion that the benefit accrued the episodic trace of HS from *repetition* was greater than the benefit accrued the episodic trace of LS. The free, category-cued, and alphabetically-cued recall and recognition of repeated items by HS and LS were compared as well as the retention of the frequency with which repeated items occurred. Second, we investigated the possibility that HS, as a matter of course, create and file episodic traces which are processed more elaborately than do LS. And, finally, we examined in detail the rehearsal habits of LS and HS in order to determine whether, during the rehearsal of items for free recall, LS and HS performed transformations that were more or less successful in assuring retention of the materials to be remembered.

VI. Repetition and IDs in Subjective Organization

The issue of how repetition operates to change retention of a memory trace is an old and fundamental one, basic to any understanding of human memory function, and it has been raised in many forms throughout the history of psychology (e.g., Ebbinghaus, 1885; Mathews & Tulving, 1973). One aspect of repetition which has been remarkably well established is the accuracy with which the human observer can report the frequency with which a repeated event occurred (e.g., Howell, 1973b; Rose & Rowe, 1976; Rowe, 1973, 1974). Furthermore, current evidence supports the notion that this is an automatic function of the memory system (Hintzman & Stern, 1978), and that not only is it unaffected by memory load or recall instructions (Flexser & Bower, 1975; Howell, 1973b), but also it is present developmentally (at adult levels of accuracy) as early as Grade Two, and it does not show improvement with practice or in response to feedback (Hasher & Chromiak, 1977). These two latter findings suggest strongly that the monitoring of frequency of occurrence of repeated events is a fundamental function of the memory system, although the actual functional utility of this activity remains a matter of speculation (Hasher & Chromiak, 1977). If one assumes, to paraphrase Lockhart, Craik, and Jacoby (1976), that the current episodic memory trace is the *outcome* of the complex analyzing and encoding operations of a pattern-recognition system whose function is to interpret incoming stimulation, then the question about the basis for frequency judgments focuses upon changes in this *outcome* as a function of there having been and/or of there now being another current representation of the same nominal event. Hintzman (1976) lists three hypotheses as to the effects upon memory traces of repetition: the trace-strength hypothesis, the multiple-trace hypothesis, and the propositional-trace hypothesis (see Howell, 1973a, for a fourth and fifth: the multiple-process hypothesis and the numerical inference hypothesis).

Although there is widespread agreement that stronger memory traces are remembered better (Goldman & Pellegrino, 1977), there are too many aspects of the composite memory which are readily accessible (e.g., aspects unique to one repetition or another) to support the notion that repetition results simply in an increment to some quantitative aspect of the strength of the representation. The multiple-trace hypothesis avoids this problem by holding that each repetition establishes a separate mental file, so to speak. In the metaphor of Ozier (1978, p. 471):

> The reader is urged to suppose that establishing an episodic trace is analogous to establishing a temporary file in some working area of the memory system. Each time a word is experienced, in a temporary file would be recorded, automatically, episodic information related to the context in which the item occurred, for example, the language of the instance, the modality in which it entered the system, and other current bits of information relevant to that episodic incident. There might even be a standard checklist of details to be recorded upon establishing a temporary file. In addition, the temporary file would contain some information concerning the semantic information of the item, that subset of information being copied, as it were, from the complete logogen located in semantic memory. There might be pointers in this file to other temporary files, and a pointer to the address of the logogen. There might be a copy of the name of the item, or a pointer to the address of the name of the work in lexical memory, or a pointer only to the logogen in semantic memory, which would allow indirect access to the name of the item when necessary.

According to this metaphor, the multiple-trace hypothesis holds that each temporary file, in addition to containing its own unique description which allows it to be accessed separately (Hintzman & Block, 1971), contains information about other occurrences of items having the same name. It is, of course, possible that both increments in strength and multiple-traces occur (the multiple-process hypothesis). The propositional-trace hypothesis (Anderson & Bower, 1974) maintains that each repetition results in the update of a single file to include the information from the current experience along with that from the past. Hintzman and Stern (1978) argue that the propositional-trace hypothesis includes the assumption that the update is voluntary, and so this explanation for how frequency effects are represented may be less attractive in the face of the evidence for automaticity of this function.

None of these accounts of the mental activities involved in frequency judgments addresses the issue of why more frequent items are better remembered. One current approach to the role of repetition *in memory* is through the study of the spacing effect (Hintzman, 1969), the persistent finding that repeated items are better remembered after distributed repetitions than after repetitions that occur in immediate succession. Hintzman (1974, 1976) has provided two extremely useful review papers in which he has suggested that the term *spacing effect* be used to refer to increase in a variety of experimental tasks over spacings of 0 to 15 sec, and the term *lag effect* (Melton, 1967, 1970) to refer to the function relating free-recall performance to the number of items intervening between repetitions during the study phase of the free-recall trial. In single-trial free recall, retention as a function of lag between repetitions increases at least up to 32 intervening items (e.g., Glenberg, 1977; Madigan, 1969), although judgments of lag are very poor (Hintzman & Block, 1973; Underwood & Malmi, 1978).

A thorough account of the various attempts to explain the spacing/lag effect is available (Hintzman, 1976); the most plausible concatenation of his detailed exposition and more current argument and counterargument is that there may be an automatic, and fundamental, central attention-distribution mechanism which directs less effort or less central processing capacity to events which repeat in immediate succession than to events which are more disparate in a sequence. There seems to be no simple answer to the more basic question as to why repeated events are better recalled than nonrepeated ones (cf. Nelson, 1977), except to suppose that implicit in the trace-strength, multiple-trace, multiple-process, and propositional-trace proposals for frequency monitoring is the notion that increments representing frequency also increase the probability that an item will be retrieved. The two experiments described below examined the relative benefits derived by LS and HS from repetitions spaced at various intervals within a series, and determined the relative accuracy with which LS and HS were able to judge the frequency with which an item had occurred.

A. THE LAG EFFECT AND IDS IN SUBJECTIVE ORGANIZATION

Experiment I in this pair of experiments replicated a study reported by Madigan (1969).[3] In that study, subjects were presented a series of 72

---

[3] Thanks are expressed to S. Madigan who shared with us his experimental materials.

words which consisted of 24 items presented once: 8 primacy items, 8 recency items, 9 midseries singles, and 24 items presented twice: four at each of lag 0, 2, 4, 8, 16, 32. There were four sets of 48 different words, and for each set was produced four different sequences of items so that the actual members of each set played different roles (repeated or not, etc.) in the different sequences. The order of sets of words and the sequence within each set were randomly assigned to each subject. The subjects pronounced each item as it appeared in the window of the memory drum; and, after the items had been presented, they were given 4 min for written free recall. In addition, beside each word recalled, the subjects wrote the frequency with which they judged that item to have occurred. (It should be noted that Madigan had shown no effect upon free-recall performance of requesting such frequency judgments. The procedure of exposure of the items and free recall was repeated four times, after which there was a final free-recall trial during which subjects were allowed 4 min for free recall of all four lists.

The results of this experiment for the repeated and the not repeated items are shown in Fig. 5 for LS and HS separately for the three rates of presentation at which the study was replicated (1.0, 2.5, and 4.0 sec per item). From the graph it is apparent that at every rate and at every class of item except recent, HS remembered better than did LS. However, all

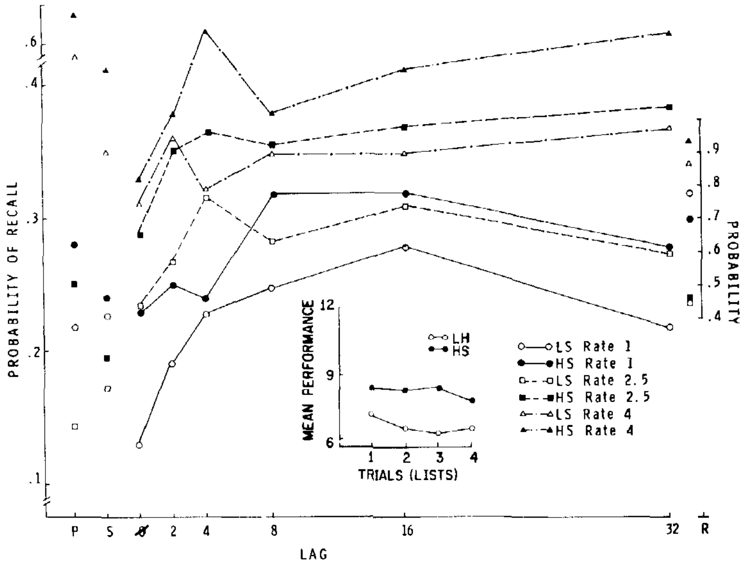


Fig. 5. Immediate free recall of LS and HS as a function of 10 trials for three rates of presentation. Also shown is mean performance for four trials (different lists).

subjects recalled greater numbers of items as the rate of presentation was increased, and the relation between lag and recall of repeated items was similar for LS and HS. The statistical analysis confirmed the results depicted here; there were significant main effects of Lags, Rates, and LS/HS.

The results for the frequency judgments are shown in Fig. 6 averaged over rates, since rates (while a significant main effect) did not interact with any other variable. The points in Fig. 6 are conditional probabilities, i.e., the probability that an item was correctly rated as to frequency, given that it had been recalled. Straight lines have been fit to these two sets of points. The graph shows that there is a slight positive slope for HS and a negative slope for LS, indicating that for LS the probability that a repeated item which was recalled was correctly identified according to frequency decreased as the separation between the repetitions increased. Figure 7 shows that both LS and HS replicated a finding reported in the original Madigan (1969) study. The finding was that the lag effect upon the free recall of repeated items held for repeated items which subjects reported as having been seen twice (2×) and not for items which subjects reported as having seen once (1×). As is obvious, LS and HS did not differ in the proportion of items recalled and misjudged as having occurred only once, but they did differ markedly in the recall of items judged to have occurred twice, although the lag effect remains prominent and similar for both groups.

The results of the free-recall posttest in which subjects were asked to

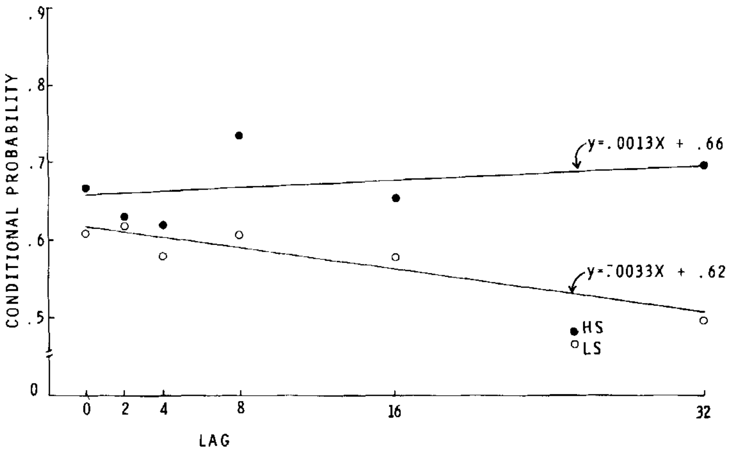


Fig. 6. Conditional probability of making a correct frequency judgment given that an item was correctly recalled.

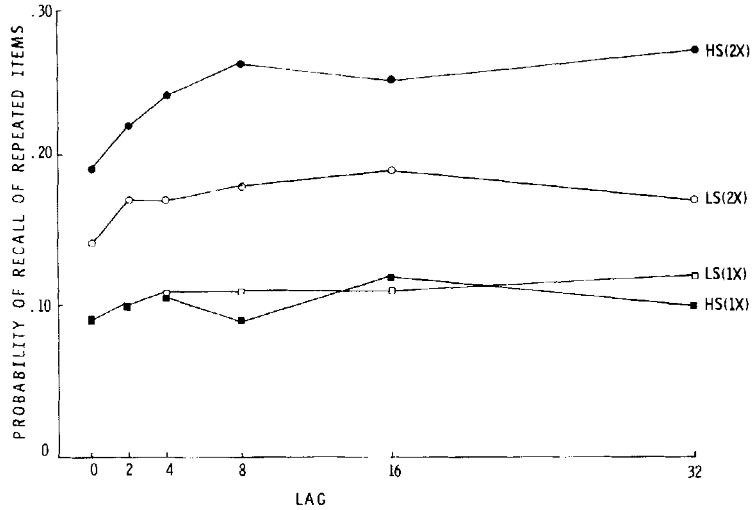


Fig. 7. The lag effect for repeated items correctly judged as having occurred twice (22) and incorrectly judged as having occurred once only (12).

recall the items from all four lists are presented in Fig. 8 for mean recall over lists and over rates as a function of lag; the interaction between lag and LS/HS was significant, $F(5,660) = 2.47$, $p < .05$. Overall recall of the four lists was 15.89 for LS and 20.89 for HS out of the 192 possible different words. While the difference is small, relatively speaking (5.0 words overall or 2.6%), it is statistically significant, $F(1,132) = 21.69$, $p < .01$, for repeated words, and also for the words presented once. Figure 8 presents a clear indication that the control of lag over posttest free recall occurred for HS but not for LS, and this may be the most intriguing result of the entire experiment. It appears as though whatever aspect of memory it is that determines the increase in retention as a function of lag, that may be shorter lasting for LS than for HS.

There seem to be two conclusions to be drawn from the present results. First, although they differ in the absolute level of correct recall, LS and HS do not differ in the shape of the function relating recall performance to lag if the recall test immediately follows the presentation of the to-be-recalled items. Whatever the correct interpretation of the lag effect, be it encoding variability of the spaced items or deficient processing of the contiguous repetitions (see Rowe & Rose, 1977), or automatic variations in attention or in central processing capacity devoted to spaced and contiguous repetitions, there seems to be a similar relative effect for both of our subjective-organizer groups. There is some evidence from Fig. 8, however, that the lag effect may be more transient for LS than for HS.

Second, it appears as though the accuracy of frequency judgments may deteriorate with lag for LS, and not for HS. We will return to these findings after Experiment II of this pair of studies is explained.

B. JUDGMENTS OF FREQUENCY AND IDS IN SUBJECTIVE ORGANIZATION

The lag experiment described above was designed, primarily, to compare the lag effect for LS and HS; the frequency judgments were made only on items that the subjects actually recalled, and there remains the possibility that LS might have been just as accurate as HS in estimating frequency of occurrence had they had the opportunity to judge all of the items they had seen during the sequence of presentations. The second study of this pair examined the frequency judgments of LS and HS in a design which was directly to compare frequency judgments of all items experienced originally.

The procedure of Experiment II (Ozier & Anderson, 1979) was as follows. LS and HS were presented one of four orders of English words in which items were presented only once or repeated 2, 4, 6, or 10 times at lags of 0, 2, 4, 8, or 16 other items. Since it has been argued strongly that "Frequency judgments depend critically upon the establishment of stable semantic encoding of repeated words" (Rowe, 1974, p. 64, see also Rowe & Rose, 1976), provision was made in the procedure for each item

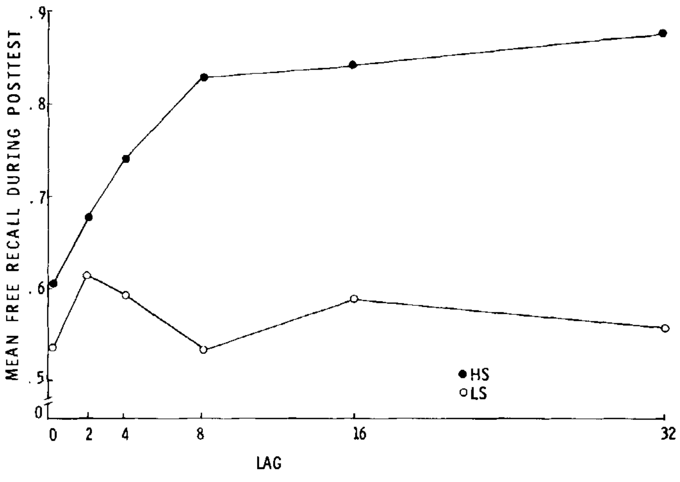


Fig. 8. The mean number of words correctly recalled as a function of lag between repetitions for LS and HS during the free-recall posttest during which subjects were asked to recall words from all four of the original lists.

to be processed semantically by requiring a semantic judgment as to whether the item presented belonged to a particular category or not. The repeated items were all positive judgments; of the single presentations, half were judgments for which the correct response was positive and half were judgments for which the correct response was negative. Each 6-sec per item trial involved the subject's reading the question, observing a slide on which a word appeared, and recording the semantic judgment by checking yes or no beside the question. Since intentional instructions have been shown to produce best accuracy for frequency judgments (Rose & Rowe, 1976), the subjects were informed in the original instructions, before the sequence of slides began, that some of the items would be repeated various numbers of times, and that after the category judgments had been completed, they would have to indicate for each word in the sequence the number of times they thought it had occurred.

There were five retention tests following the sequence of category judgments: frequency recall, free recall, recognition, category-cued recall, and alphabetic-cued recall. For the frequency judgments, the subjects were given a randomly ordered list of the items they had seen, and 4 min to rate the frequency with which each item had occurred on a scale which ranged from 1 to 10 items. The free-recall test was next. Subjects were asked to write as many of the words they had seen on the screen in any order they liked. Following this, subjects were given a test of recognition: 5 min to check from a list of 160 items, those 80 which had been presented during the initial stage of the experiment. The category-cued recall test provided the subjects with 5 min to write the items they recalled beside the names of the 50 categories to which they belonged. Finally, the subjects were given a random sequence of the letters of the alphabet with which the items had begun, and for 5 min they were asked to try to remember the words according to the sequence of initial letters provided.

For the purpose of clarity of exposition, the results of this experiment will be presented somewhat out of order, first the results of the recall tests and then the results of the frequency judgments. The recall tests (free, alphabetic-cued, and category-cued) are shown in Table II for the total number of items correctly recalled and the repeated items separately. The recognition test results indicate that a trace of some sort had been established for almost every item in the list repeated or not; the probability of recognition was quite high (.84) and for repeated items it was even better (.94). Although there was a significant difference in overall recognition performance between LS and HS, the LS did reach 82% accuracy of recognition even on the overall measure. Those traces reflected in recognition were somewhat less accessible when retrieval according to the name of the category to which they belonged was required; the average

TABLE II

PROBABILITY OF CORRECT RECALL OR RECOGNITION BY LS AND HS FOR ALL ITEMS AND FOR REPEATED ITEMS SEPARATELY[a]

	All items		Repeated items	
Test	LS	HS	LS	HS
Recognition	.82	.86	.93	.95
Category-cued	.57	.65	.85	.91
Alphabetic-cued	.28	.38	.43	.57
Free recall	.28	.38	.48	.58

[a] In each case except for the recognition of repeated items the difference between LS and HS was significant at better than the 5% level of confidence.

probability of recall was .61 overall and .88 for repeated items. Finally, the memory traces established during the semantic-judgments task were rather difficult to retrieve according to the alphabetic initial-letter cues ($p = .33$) and free recall ($p = .33$), although in both of these cases HS performed at least 10% better than LS (see Fig. 9). In summary, the data from the recognition tests indicated there had been a memory trace for almost every item, that over half of those traces were retrievable in response to the name of the category to which they belonged, and that only one out of three was retrievable in response to the initial letter of the item or with no cue whatsoever. However, in every case but one, LS recalled fewer items than HS.

Having attended to the recall characteristics of the episodic traces established here (cf. Ozier, 1978), we may turn to the consideration of the indices of frequency attendent upon those traces. Overall, the average judgments of frequencies of 2, 4, 6, and 10 presentations were fairly accurate, 2.4, 5.2, 7.3, and 9.2 to be exact. However, HS judgments (2.3, 4.8, 6.9, and 9.2) were closer to the actual frequencies than were LS judgments (2.5, 5.6, 7.6, and 9.3). The same result occurred with the items which had occurred once. The singly occurring items were rated as having occurred 1.26 times; those which had had positive semantic judgments were rated as having occurred more frequently than those which had had negative semantic judgments (1.31 vs 1.21, $p < .01$), but HS were more accurate in their ratings of singly occurring items than LS (1.13 vs 1.39, $p < .05$). Figure 10 shows in more detail, the average frequency judgments as a function of lag between repeated items, for LS

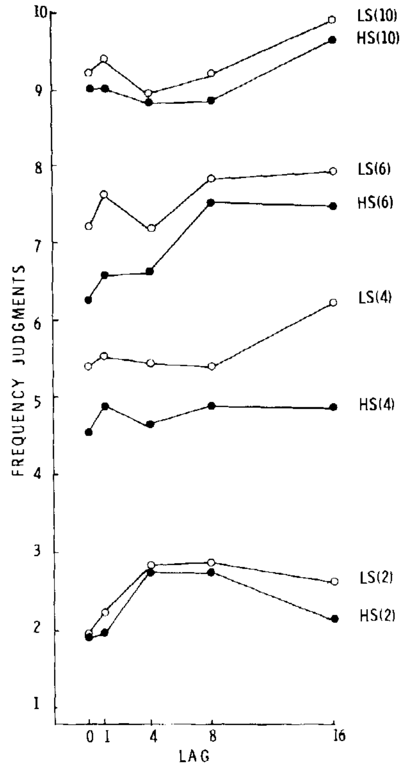


Fig. 9. Free recall as a function of the frequency with which items were repeated during the semantic judgments task.

and HS and for each repetition frequency separately. The graph reveals that the shape of the relation between lag and frequency judgments varied somewhat for the four actual frequencies, and analysis of variance confirmed that this was so, there being a significant interaction between Frequency and Lag, $F(12,552) = 1.94$, $p < .05$. For both LS and HS, the more frequently an item was repeated, the greater was the overestimation at long lags. This interaction between Frequency and Spacing/Lag has been reported at least three times before (Rose & Rowe, 1976, Experiment I, Experiment II; Rowe & Rose, 1977), and it has been interpreted as enhancement of the effect of spacing by higher frequencies of occurrence within a series.

With regard to the main variable of interest here, however, the difference between LS and HS in the judgment of frequency, at least two points can be made. First, the enhancement of the effect of spacing between repetitions by higher presentation frequencies occurred for both low and high organizers; there was not a significant interaction between all three variables, $F(12,552) < 1$. Second, although the difference between LS and HS appears to have varied with the presentation frequency in the analysis of variance, the F-value for the interaction between Frequency and LS/HS was just short of significance, $F(3,138) = 2.63$, and there was a significant main effect of LS vs HS, $F(1,46) = 4.56$, $p < .05$. A strict interpretation of these data would require the conclusion that overall, regardless of the impression provided by the graph, LS gave higher judgments of the frequency of with which the repeated items had occurred than did HS.

The results of this pair of experiments, Experiment I focusing upon the effect upon the immediate free recall of LS and HS of repeating items at different intervals, and Experiment II focusing upon the effect upon frequency judgments of LS and HS of repeating items at different intervals, provide compelling evidence that whatever may be the correct explanation of the spacing/lag effect, it exerts its control for both LS and HS in a similar manner. An additional similarity between LS and HS was the degree to which repetition improved their performance in recognition, category-cued, alphabetic-cued, and free recall of Experiment II. Although the change in probability of recall from one repetition to ten

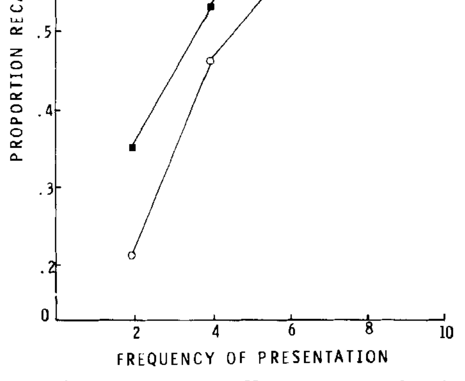


Fig. 10. Mean frequency judgments for LS and HS as a function of the lag at which the repeated items were repeated for items 2, 4, 6, and 10 times.

repetitions for these tasks varied (an improvement of 3% for recognition, 11% for category-cued, 11% for alphabetic-cued, and 45% for free recall, the data for recognition and category-cued recall being subject to a severe ceiling effect), in each case there was a similar improvement for both LS and HS as a function of the repetition variable (although in absolute terms, LS recalled fewer items, of course).

These data are offered here as some support for the more general notion that the representation of frequency attendent upon the episodic trace of an item in memory functions in the retrieval of that item (cf. Tversky & Kahneman, 1973). This argument has been made previously for recognition memory (e.g., Underwood, 1972), and it is made here for other memory tasks as well. There is a difference of opinion about whether frequency judgments reveal anything at all about recall. For example, Rowe (1973) has reported that the recognition confidence ratings of homonyms was not affected by whether they were repeated in the same or in different contexts but frequency judgments were; Howell (1973b) has reported that frequency judgments were independent of instructions as to whether the items would have to be judged or not, whereas recall was improved with relevant instruction. On the other hand, Rowe (1974) has shown that frequency judgments vary with the level of processing (counting the number of syllables vs judgments of connotative meaning) as does free recall. It seems as though the matter could be argued either way: (*a*) frequency judgments are based upon an aspect or attribute of the memory trace not involved in retrieval or (*b*) frequency judgments and retrieval are a function of the same outcome of perceptual analysis. The fact that LS and HS differences occur in every retrieval task employed in these experiments and also in frequency judgments suggests that (*b*) might be the correct interpretation.

The interpretation of the main findings of this pair of experiments, that frequency judgments are *higher* (less accurate) for LS and free recall *lower* for these same subjects, may lie in the variability of the encodings established by LS and HS for the same repetitions of the same items. Although inflated frequency judgments have been reported as a result of a number of experimental manipulations, most of these were addressed in the design of our experiments. For example, Begg (1974) has reported that terminal tests produced higher judgments than did continuous in series tests, but all of our subjects did terminal tests. There is a difference between judgments based upon the level at which the items are processed (Rowe, 1973), but our subjects were all required to do semantic processing for every item. Frequency judgments after incidental instructions are higher than they are after intentional or memory instructions (Rose & Rowe, 1976), but all of our subjects were told (in no uncertain terms) that they would have to judge each item for frequency of occurrence after the sequence of presentations had been terminated.

What was not controlled in our studies was the variability of the semantic encoding engaged in by LS and HS. The odd finding that *different* encodings of repeated items produce *lower* frequency estimates has been reported where the difference in the encodings was controlled by varying the phrase context in which homonyms were presented to result in either same meaning or different meaning codes (Rowe, 1973), or the difference was defined as judgments on semantic scales which varied or sentence contexts which varied in content (Hintzman & Stern, 1978), or different filler words preceded and followed the repeated items (Hintzman, 1974, p. 94). The results of these studies have shown higher frequency judgments for consistent encodings and lower frequency judgments for variable encodings. We suggest that HS may engage in more variable semantic encoding than LS, this resulting in consistently lower frequency judgments.

With regard to the difficult question as to why variable encodings should lead to *lower* frequency judgments but *higher* retention performance, there is valuable theoretic support (Bower, 1972; Martin, 1968; Melton, 1970) but mixed evidence. Where encoding variability has produced better retention (Bevan, Dukes, & Avant, 1966; Hintzman & Stern, 1978), it has been under circumstances in which the context of the item has changed with repetition, but not the meaningful representation assumedly aroused by the experience of that item (see also Rowe, 1973, Fig. 3, where the difference is in this direction although not significant). The conclusion offered here with regard to the results of our two experiments is that HS may be more successful in recording a repetition of an item as having the same meaning but a different context (temporal position in the sequency vis-à-vis the other items for example), or at the same memory node, but with different list markers (à la Anderson & Bower, 1972). Recording a repetition of the same logo within a different context of spatiotemporal attributes (Underwood, 1969) might result in lower frequency judgments and higher retention scores.

Another possible explanation would incorporate Jacoby's (1974) important conclusion concerning implicit mental contiguity in memory. Jacoby maintained that in order for one member of a category to become a successful cue for another member of the same category, not only did the traces of the two items have to be made contiguous in memory, but also the traces of the related items had to change, presumably that change in the direction of adding to the retrieval information of each trace the instruction that the name of the trace would be brought to consciousness when the other category member was experienced, or recording of the ''address'' of the related item, etc. In the present context, then, it may be that with repeated experiences, of the same nominal item, HS are more successful in recording with each new episodic file, the location of other files of the same nominal item, or indeed, the number of those currently in existence.

The immediately preceding paragraphs are entirely speculative, but they do present notions concerning the basic underlying processing differences between LS and HS which are testable. The experiments are not done, but it does make some sense to suggest that success in recording contextual episodic information might underly not only the differences in frequency judgments and retention performance found in the present studies, but may in fact underly the difference in performance in multitrial free recall and the subjective organization found therein (see Runquist & Runquist, 1978, for paired associate learning). It may be far fetched to imagine that better subjective organization results from more successful or more complete implicit mental contiguity, i.e., the recording on the episodic files of those items the addresses of the members of a cluster, group, or sensible (to the subject) sequence. If we consider the repetition of an item in a single long sequence to be somewhat analogous to the repetition of the same item from trial to trial in multitrial free recall, then the leap from frequency judgments to subjective o-ganization may seem less hazardous. In any case, the proposal is made here that IDs in subjective organization may, when understood, reveal a great deal about the basic functioning of the episodic memory trace and its control processes

It may be a matter of some curiosity to the reader to know at this point whether the differences between LS and HS retention performance disappear or not when the perceptual analysis of incoming stimuli is directed toward the physical or phonemic (rather than the semantic) characteristics of the items to be recalled. It is to the investigation of LS/HS differences at these various levels of processing (Craik & Lockhart, 1972; Hyde & Jenkins, 1973) or after various perceptual analyses to which we now turn our attention.

VII. ''Different'' Encoding Operations and IDs in Subjective Organization

The experiment reported next (Ozier & Sperry, 1978) was designed to compare free recall and recognition of LS and HS of items processed to various ''levels'' in accordance with the metaphor for memory offered in

1972 by Craik and Lockhart [although the phrase "different[4] processing" (Nelson, 1977, p. 168) is probably more correct]. While the current discussion over the theoretic value of the level of processing framework continues (Cermak & Craik, 1979), and predictions made by the various versions of this metaphor get tested over and over again (Baddeley, 1978; Eysenck, 1978; Nelson, 1977; Postman & Kruesi, 1977), the positive influence of the original notion that memories are the outcomes of various perceptual analyses (Craik & Lockhart, 1972), i.e., that "what the subject was instructed to do with the input was the most important determiner of the nature of what was remembered" (Jenkins, 1977, p. 426) is unquestionable. It is not clear that the specifics of the original formulation are correct [consider the intriguing alternative that memories are remembrances of the analytic operations themselves (Kolers, 1979)], nor do the current addenda or revisions involving distinctiveness of traces and elaboration of encodings seem particularly attractive changes at this time, and the picture becomes even more complex when one considers Folkard's (1979) report that subjects spontaneously adopt a physical bias in the morning and a semantic one in the evening. How these spontaneous cyclic changes in information-processing style might interact with subjects' judgments and various retrieval tasks boggles the mind. Nevertheless, the procedure of asking subjects to make incidental judgments about verbal items and then examining the traces of those items by requiring retention tests of various sorts has been an extremely useful one, not only because of closer control over the mental activities engaged in by subjects during their initial experience with the material to be remembered (e.g., Schulman, 1974), but also because the interaction between encoding operation and retrieval requirements has become much more amenable to careful examination (Bransford, Franks, Morris, & Stein, 1977).

The experiment reported below was modeled after the basic paradigm of Craik and Tulving (1975). Our purpose was to determine whether LS and HS differences in free recall and recognition memory could be altered by having subjects make different kinds of judgments concerning the items they would have to recall and recognize. The different kinds of judgments being assumed to require different perceptual analysis and so having different perceptual outcomes, and so leaving different qualities of memory traces, it was expected that more detail concerning the episodic

---

[4] While we bow to Nelson's strong arguments, we do so with some aesthetic reluctance. As a member of "A Chorus Line" puts it: "*Different* is nice, but it sure isn't *pretty*." The metaphor of "Levels" seems to have a very strong psychological attraction despite the weakness of the specific predictive power shown to date.

traces of LS and HS would be revealed. A specific impetus for the study was Nelson's (1977) report that taxonomic organization had occurred after semantic processing but not after phonemic processing. Our expectation was that LS and HS would differ only on items encoded semantically.

The three kinds of judgments we selected were representative of three classes of encoded tasks: judgments of the appearance of the stimulus (type case), judgments concerning the sound of the word (as to rhyme quality), and judgments concerning the meaningful aspect of the word (appropriateness in a particular sentence context). Despite great sympathy for Lockhart's (1979) strong plea that "orienting tasks should never be selected to represent different levels of processing. They should be chosen on the grounds that they model (that is, capture the essential features of) an aspect of cognitive processing that is functionally important and that they have a high degree of ecological validity" (p. 79), we must confess that when this experiment was designed, the tasks were selected because they had been explored extensively, they had been shown to produce very robust differences, and it was intended that the previously published data provide a basis for comparison of LS and HS interactions. Notwithstanding this admission, however, we believe it is possible to justify our choice on other grounds. Certainly it seems almost universally accepted that the appearance, the sound, and the meanings of a verbal item require perceptual analysis outcomes which must differ from one another in some ways; as for functional importance, the case in which an item appears guides our analysis of printed language and its interpretation in terms of surface structure as well as the IMPORTANCE of some items over others, and the understanding of the abstraction of meaning is central and critical to any theory of mind. The judgment of rhyme does seem a little more frivolous in terms of the ecological validity of the orienting tasks we chose; nevertheless, rhyme is a quite fascinating aspect of the sound of a word, one which the child can appreciate and produce as early as 5 years of age (Jusczyk, 1977), and so may well be a key to understanding the child's (and the adult's) awareness of the phonological properties of the language (which may indeed be a necessary prerequisite for reading readiness, see Rozin & Gleitman, 1977).

The experimental details were as follows. The materials used were sets of 36 singular nouns, 3 to 7 letters in length, of medium to high frequency of occurrence according to Thorndike and Lorge (1944). The subjects were shown each of the 36 target items twice each, on the screen of an oscilloscope slaved to a PDP-12 computer. Each presentation was preceded by a question regarding the target item to follow. There were equal numbers of questions regarding the case, rhyme, and appropriateness in a sentence context, and equal numbers of each required positive and negative judgments. The two rhyme questions for the same item contained different comparison words; the two sentences for the semantic judgments contained different sentence contexts but ones which required the same meaning of the target item to be used. The sequence of presentations was such that all items were tested once before they were tested again, and equal numbers of subjects answered one of the six questions possible (case/positive, case/negative, rhyme/positive, rhyme/negative, sentence/positive, sentence/negative) for each item. Each trial consisted of a 5-sec interval in which the question appeared on the screen after which the target item was shown. The response was given by two response keys which the subjects pressed with the index fingers to indicate a yes or a no answer. Half of the subjects used their preferred hand for yes and nonpreferred for no and half vise versa. The computer recorded the time from the onset of the target item to the depression of one or another of the response keys (RT).

The initial instructions to the subjects indicated that this was an experiment about the speed of perceptual judgments. After the judgments had been made, the subjects were required to recall as many of the words as they could remember, in any order they liked (for 4 min). Also, subjects were given a recognition test in which they were required to check those items which had been seen on the screen out of a randomly arranged series of 108 items, the 36 target items and 2 distractors matched with each target for frequency of occurrence.

The LS and HS did not differ on the number of errors they made on the original judgments (the probability of an error was .051 and .047 for LS and HS, respectively). The reaction times to respond yes or no to the various judgments are shown in Fig. 11. The average time taken for case judgments was 964.9 msec, for rhyme judgments 1025.1 msec, and for sentence judgments 1075.0 msec. The graph shows positive responses were made more quickly for case and sentence judgments, but less quickly for rhyme judgments than were negative judgments, a finding reported also by Craik and Tulving (1975, see their Fig. 1). On the average, for both positive and negative responses, and for every type of judgment, HS performed more quickly than did LS, the average difference in their RTs being 77 msec, but these differences ranged for the different types of judgments, being 145.0, 25.3, and 60.8 msec for case, rhyme, and sentence judgments, respectively. The analysis of variance confirmed the information in the graph shown here; there was a significant interaction between C/R/S and LS/HS, $F(2,84) = 3.18$, $p < .05$, and a significant interaction between C/R/S and Yes/No, $F(2,84) = 7.02$, $p < .01$. It appears as though HS are faster than LS in analyzing printed

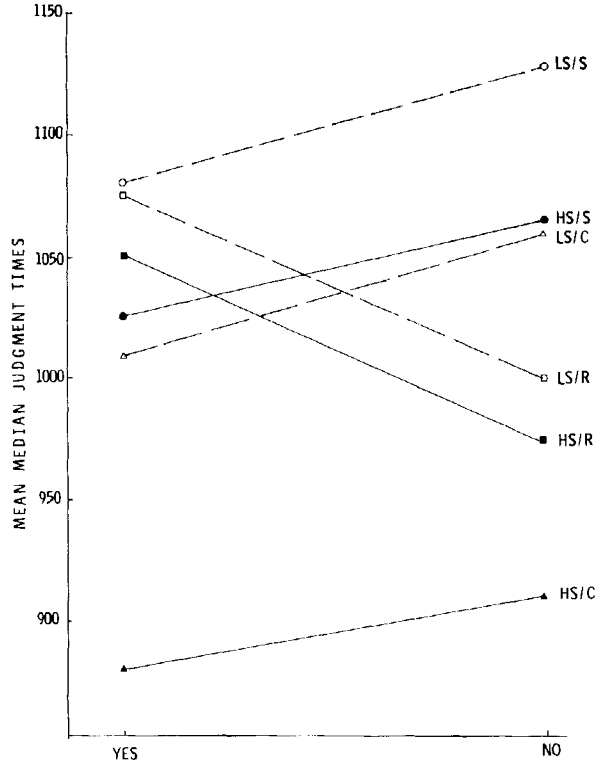


Fig. 11. Mean median judgment times for positive and negative judgments made by LS and HS for case, rhyme, and sentence questions.

verbal material, particularly in extracting information related to the appearance of the stimulus items.

Performance on the free recall task was relatively poor, an average of 7.75 words for LS and 10.19 words for HS out of the 36 possible. Figure 12 shows the probability of correct free recall of the target items by LS and HS as a function of the different independent variables in this experiment. The figure shows that more items were recalled of those given positive judgments (10.74) than were given negative judgments (7.21) (see Schulman's 1974 discussion of congruous and incongruous judgments), and that case, rhyme, and sentence judgments varied as well (averages of 3.02, 4.67, and 10.25 items, respectively). A similar result is depicted in Fig. 13 where the results of the recognition test are shown. Analyses of variance for recall and recognition separately revealed significant interactions ($p < .01$) between C/R/S and Yes/No and main effects of C/R/S, Yes/No, and LS/HS, the statistics for free recall being $F(2,84)$

= 4.55 for the interaction, $F(2,84) = 97.8$, $F(1,42) = 27.8$, and 8.66 for the main effects in order. The source of the interaction in both cases seems to have been a smaller difference between positive and negative case judgments in both recall and recognition.

What can we say about IDs in subjective organization from this experiment? The first point seems to be that the time taken to make a judgment does not seem to be related, particularly, to the probability that the trace established in making the judgment will be retrieved. HS spent less time in analysis of the incoming stimulation, made as many correct judgments, and more correct recall and recognition responses than did LS. This finding that reaction time does not predict the probability of recall confirms a conclusion drawn by Craik and Tulving (1975) who found that an easy but deep task (sentence judgments) resulted in better retention than a complex but shallow one (vowel, consonant pattern judgments); the interpretation of their result was that it is not because they take more time that semantic judgments produce better retention.

The most dramatic demonstration of this in our own data is the comparison between LS recall of case targets and HS recall of sentence targets. Figure 10 shows that these judgments took almost the same amount of time (an average difference of 7.2 msec), and yet the porbability of recall of case targets by LS was .10, whereas the probability of recall of semantic targets by HS was .48. Certainly the length of time taken to make a judgment is not the critical factor in the successful retention of memory trace; what seems to matter is first, what is done during the time the judgment is being made, and second, who does the

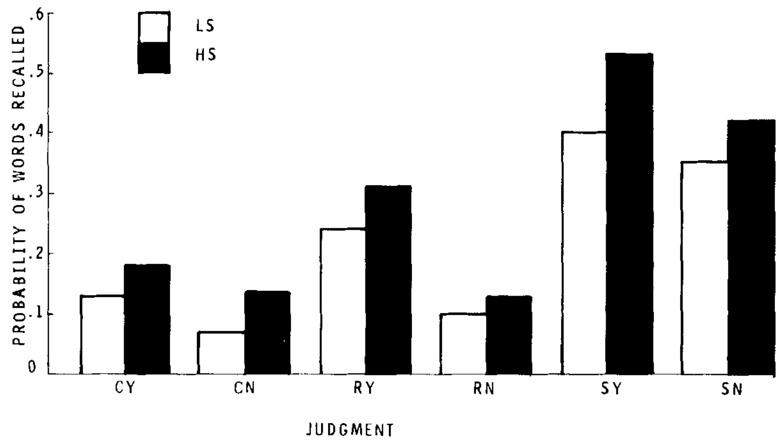


Fig. 12. Free recall as a function of positive (Y) and negative (N) responses for Case (C), Rhyme (R), and Sentence (S) judgments.

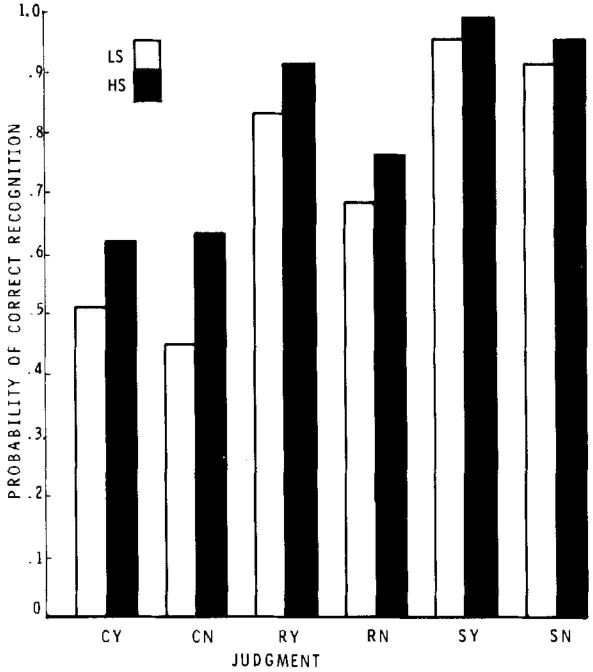


Fig. 13. Recognition as a function of positive (Y) and negative (N) responses for Case (C), Rhyme (R), and Sentence (S) judgments.

doing. The data of the present experiment seem to require the conclusion that whatever HS are doing which results in better remembering, that process is involved in the establishment in the outcomes of at least three types of perceptual analysis; the further discussion of these persistent differences will be delayed until the final discussion when we will attempt to integrate the message we believe they convey.

VIII. Rehearsal and IDs in Subjective Organization

The last two sets of experiments presented in this article report the free recall and recognition of LS and HS after two different types of repetition experience, i.e., after two different types of rehearsal activities. The distinction between types of rehearsal has been enunciated in several different ways, not always compatible with one another. For Craik and Lockhart (1972) Type I rehearsal was said to be repetition of analyses which had already been performed, whereas Type II rehearsal resulted in more extensive (deeper) analysis of the stimulus material. Seemingly, Type I rehearsal was functionally responsible for maintaining information for the current time period; Type II rehearsal contributed to the establishment of a trace which might be addressed at a more remote instant. For Rundus (1977), Maintenance Rehearsal (Type I) had two defining characteristics: (*a*) the subjects' intent to preserve information in an active state, and (*b*) no relation between the amount of maintenance and the probability and the retention of the item material being repeated; Coding Rehearsal resulted when there was an intent to retain the material being repeated.

Since the data concerning whether Type I rehearsal results in a change in the trace of the material to be remembered are mixed (Nelson, 1977), there being strong evidence that recognition improves with Type I rehearsal (Glenberg, Smith, & Green, 1877; Woodward, Bjork, & Jongeward, 1973) and little evidence that Type I rehearsal has any effect upon free recall, Type I rehearsal has been further differentiated. Woodward *et al.* (1973) have suggested that any rehearsal (rote or not) may result in transfer of an item to long-term memory, where the encoding of that item may contribute to its recognition, but that an improvement in long-term recall will result only if constructive, associative rehearsal takes place. Glenberg *et al.* (1977) have made a simpler distinction between two qualities of Type I rehearsal, Maintenance Rehearsal and Primary Rehearsal, the former being repetition which has no memorial aftereffects, and the latter being memorial effects which allow better recognition but not better recall. However, there does not seem to be any good evidence that Maintenance Rehearsal having no memorial aftereffects exists at all (Nelson, 1977), and so, for the time being at least, Primary Rehearsal and Type I rehearsal seem to be synonymous terms.

A. TYPE I (PRIMARY) REHEARSAL AND IDS IN SUBJECTIVE ORGANIZATION

The purpose of the first study reported below was to examine the recognition and recall of LS and HS after Type I rehearsal. The experiment was essentially a replication of Glenberg and Adams (1978), in which a clever task was combined with a sensitive recognition test (one which was designed after Coltheart, 1977) to reveal the semantic and acoustic characteristics of the traces established during the repetition of the material to be tested.[5] The task combining a concurrent activity (digit recall) with the rehearsal of other material (word pairs) was one for which Glenberg *et al.* could show that cognitive capacity was shared by the two

---

[5] The generosity of A. Glenberg, who provided us with 72 rehearsal pairs and the four distractors for each target word, is gratefully acknowledged.

activities, but the portion of cognitive capacity was minimal for the rehearsed word pairs.

The procedure was as follows. LS and HS subjects participated individually in an experimental session which consisted of 81 trials on a delayed digit-recall task. During each trial, four digits were selected randomly (as programmed onto a PDP-12 computer) for display (by the PDP-12 computer) on the screen of a slave oscilloscope for 2 sec, followed by a pair of words for 2 sec, followed by a blank screen which remained for enough time to allow for 1, 5, or 10 overt repetitions of the word pair. The computer controlled a tone which repeated during the interval and set the pace for the subject's rehearsals. Each trial ended with a 5-sec interval for digit recall, and a 2-sec warning that the next quartet of digits was about to appear. Of the 81 trials of this sort, 3 were preliminary practice, 3 were primacy, and 3 recency throwaways, and the 72 remaining trials constituted the data proper. After the last trial, subjects were given a five alternative forced-choice recognition test.

On the recognition test there were 78 lines; on each line there were five words, one of which had been one of the two words of the pair rehearsed. The four distractors were (*a*) a synonym of the target item (S for semantic distractor), (*b*) a rhyme of the target item (A for acoustic distractor), (*c*) a control synonym (CS), and (*d*) a control rhyme (CA) (cf. Coltheart, 1977). The purpose of the control distractors was to reveal any systematic bias of subjects toward making semantic or acoustic errors. An example of one five-choice group of one of the set of five from which the subjects' were required to choose the target item (T) is: STALL, CHILD, HALL, YOUTH, BOOTH, where YOUTH was the target (T), CHILD the semantic distractor (S), BOOTH the acoustic distractor (A), STALL (similar in meaning to BOOTH) the control semantic distractor (CS), and HALL (a rhyme for STALL) the control acoustic distractor (CA). The relation of the distractors to the target and to each other, the position of the targets and the distractors in the series was counterbalanced over three blocks of 24 lines. Following the recognition test, one member of each pair of rehearsed items was provided as a cue for its partner, and the subject's task was to write in the co-rehearsed word.

The results of the delayed digit-recall test have been reported already (Section IV). On the average, 83.4% of the digit quartets were correctly recalled, and LS did not differ significantly from HS on this task. The results of the test for recognition of the target items are shown in Fig. 14 as a function of the probability of correct or incorrect choice. On the average, LS selected 21.6 of the 72 target items, and HS selected 28.8, a difference of 7.2 which was statistically significant, $F(1,28) = 8.44$, $p < .01$. The inset in the graph shows recognition of the targets as a func-

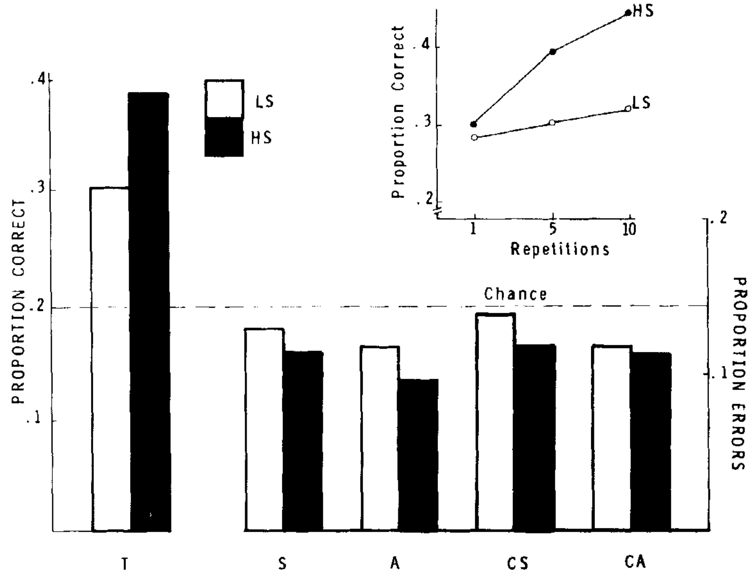


Fig. 14. Recognition after Type I Rehearsal for target (T) items and for types of distractors: Semantic (S), Acoustic (A), Semantic Control (SC), and Acoustic Control (AC). Shown also is correct recognition as a function of the number of rehearsals made.

tion of the number of times the word pair had been rehearsed. The effect of the number of repetitions was statistically significant, $F(2,56) = 6.16$, $p < .01$, but the interaction was not, $F(2,56) = 2.01$, $p > .05$.

A search in these data for the changes in the traces established by Type I rehearsal which would account for the obvious improvement in recognition as a function of the number of repetitions supported the Glenberg and Adams (1978) argument that the acoustic representations of the items were strengthened as repetitions increased. Table III shows that the probability of selecting an acoustic distractor of the target item decreased as repetitions increased for LS and HS separately, and for the two combined, and two of the analyses we attempted did show a significant decrease in the selection of the distractor related acoustically to the target items as repetitions increased. In one of those analyses,[6] the ratio of T to T + A was calculated for each subject for each level of the repetition variable, and those ratios were entered into an analysis of variance for the effects of LS/HS and Repetitions. There was a significant main effect of (1, 5, and 10) Repetitions, $F(2,56) = 3.47$, $p < .05$, the means for these

---

[6] John Barresi suggested these analyses.

TABLE III

THE PROBABILITY WITH WHICH THE TARGET ITEMS AND FOUR KINDS OF DISTRACTORS WERE CHOSEN BY LS AND HS SEPARATELY AND COMBINED

Number of rehearsals	Target (correct)	Semantic distractor	Acoustic distractor	Semantic control	Acoustic control
			LS		
1	.288	.168	.177	.196	.156
5	.303	.182	.159	.194	.165
10	.323	.188	.154	.193	.170
			HS		
1	.323	.173	.145	.158	.181
5	.397	.145	.139	.175	.149
10	.449	.153	.118	.160	.139
			Combined		
1	.305	.170	.161	.177	.169
5	.350	.163	.149	.184	.157
10	.386	.171	.136	.178	.154

being .65, .70, and .74. None of the analyses involving semantic distractors showed significant effects.

The data from the cued-recall test were extremely easy to score. Of the 72 pairs of items which the subjects had rehearsed, one of each of which was provided as a cue for retrieval, LS subjects remembered a total of 4 correctly with the co-rehearsed word and 8 correctly written anywhere on the recall test; HS subjects remembered 4 and 6. These results finally permit us to lay down the associative hypothesis raised very early in this article (Section V). It is obvious that HS have no magic association network for after repeating many of these pairs of words 10 times over, they were completely unable to retrieve one member of a pair given the other.

An old-timer colleague has argued, however, that the cued-recall test designed by Glenberg *et al.* was not appropriate to testing the automatic establishment of associations. This might better be examined by the use of a modified free-recall test originally designed by McGovern (1964) to examine associations in the absence of the requirement for recall of the responses of the paired-associates after paired-associate learning. The subject would be required to join the members of the co-rehearsed pairs of words provided in scrambled sequence. If the LS and HS do not differ on this task, then the associative hypothesis could be dismissed finally. The data for this comparison will be collected shortly. To sum up then, the data from the digit-recall test indicated that HS and LS did not differ in short-term retention with minimal cognitive capacity required for the rehearsal activity, but the recognition of the target items was better by HS than by LS by quite a wide margin. The implications of this outcome are straightforward. Whatever distinguishes HS and LS during free-recall memorization may also serve the HS during Type I rehearsal, i.e., the difference occurs even when there is no intent to memorize, there is no trace for recall, and there is minimal cognitive capacity devoted to the rehearsal activity.

B. ELABORATIVE REHEARSAL AND IDS IN SUBJECTIVE ORGANIZATION

This last section describes our ongoing attempts to produce HS performance in LS by giving differential instructions. The research began some years ago with a miniexperiment in which rehearsal patterns of LS and HS were examined according to the overt rehearsal task introduced by Rundus and Atkinson (1970), and further explored since by several others (including Brodie & Prytulak, 1975; Cuvo, 1975; Horton, 1976; Murdock & Metcalfe, 1978; Nelson, 1977; Rundus, 1971; Woodward *et al.*, 1973). In our original study, LS and HS were asked to memorize a 20-item list (in which each word began with a different letter of the alphabet and belonged to a different taxonomic category) for four trials. Each trial consisted of the presentation of the words one at a time in a different order; the duration of the presentation of each word was controlled by the subject who had an electronic, two-button, stop-start box by which he/she controlled the advancement of a slide projector. Word presentation was, then, completely self-paced. For each trial, after all 20 words had been presented, there was a 90-sec period for written free recall. Subjects were given instructions to rehearse out loud in any way that would help their recall of the words and to try to remember as many words as possible.

The results of this study were analyzed in a number of ways. We examined the time taken for rehearsal of each item as a function of the serial position in which it occurred, the number of repetitions of each item, the size of the rehearsal set at every serial position (i.e., the number of different items included in the rehearsal period at each serial position), and the probability of recall. Presented in Fig. 15 are the rehearsal times per serial position. Consider the curves for LS and HS. It is perfectly apparent that LS and HS differed dramatically in the manner in which they devoted time to the rehearsal of the items to be remembered during

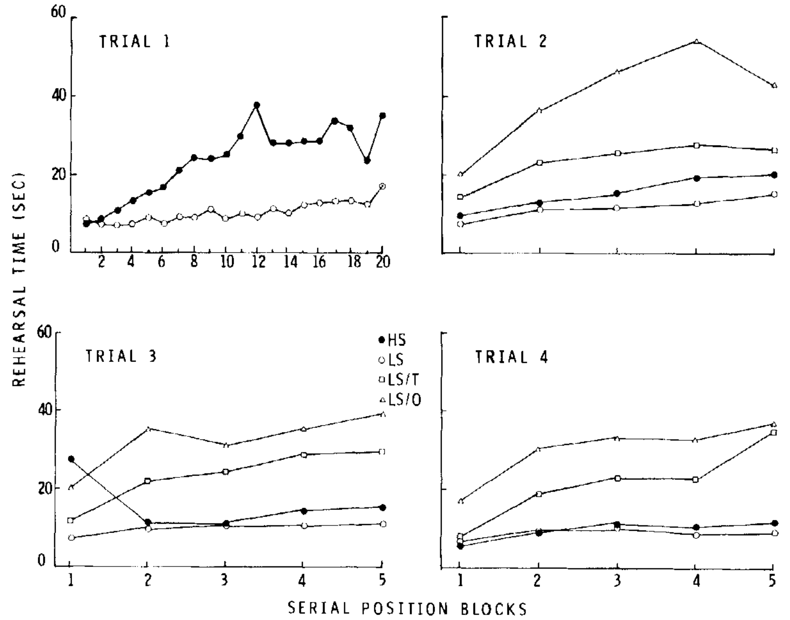


Fig. 15. Rehearsal times as a function of blocks of serial positions (blocks of two for Trail 1 and blocks of four for Trials 2, 3, and 4) for HS and three sets of instructions for LS.

Trial 1. After that, there appear to be no differences, and this combination of results was reflected in the interaction between LS/HS and Trials, $F(3,54) = 5.25, p < .01$.

Since it was quite clear that any comparison of LS and HS would have to be qualified by the Trial 1 difference in Fig. 15 (any of the other measures of rehearsal listed above), we selected three more groups of LS and performed the following variations upon the original theme. First, *for all of these groups,* we had the times for rehearsal at each serial position during Trial 1 controlled automatically instead of self-paced by LS, those times being closely yoked to the pattern shown by HS. Item 1 was presented for 10 sec; then each item was presented for 2 sec more than the previous item up to item 11; thereafter there were 30 sec given each word. For Trials 2, 3, and 4, rehearsal was self-paced by the subject as it had been for the original subjects. In one condition (LS/T for Time), the same instructions applied as had for the original LS group. LS/T differed from LS only in that they were required to rehearse, during Trial 1, just as long per serial position as had HS. The subjects in a second condition (LS/R for Repetition) were told to repeat the items as frequently as possible during the time provided for rehearsal. The performance of these subjects was indistinguishable in all ways from that of LS/T, and so the characteristics of this group will not be presented in detail. In the third condition (LS/O for Order) subjects were instructed to try to rehearse the words in the same order during each rehearsal interval.

The times taken to rehearse during Trials 2, 3, and 4 when they were free to pace themselves, by LS/T and LS/O, are to be found by returning to Fig. 15. LS/O took more time after Trial 1 than did LS/T, and both took more time than did HS and LS which did not differ from one another. It seemed as though our manipulation had been successful at least to the extent that LS were now using the opportunity provided for rehearsal.

The next figures show what LS/T and LS/O used the rehearsal period for. Figure 16 shows the number of repetitions or the number of times an item was repeated over the entire trial as a function of the serial position of the item in the presentation sequence for HS, LS, LS/T, and LS/O. The striking aspect of this graph is that on Trial 1, HS rehearsed each word more times than LS; but, after that, HS, LS/T, and LS groups did not differ terribly in the average number of repetitions made. However, LS/O continued to repeat each item many more times. The downward trend of these data reflects the fact that items presented last in the sequence had, of course, fewer opportunities for rehearsal than did items early in the se-

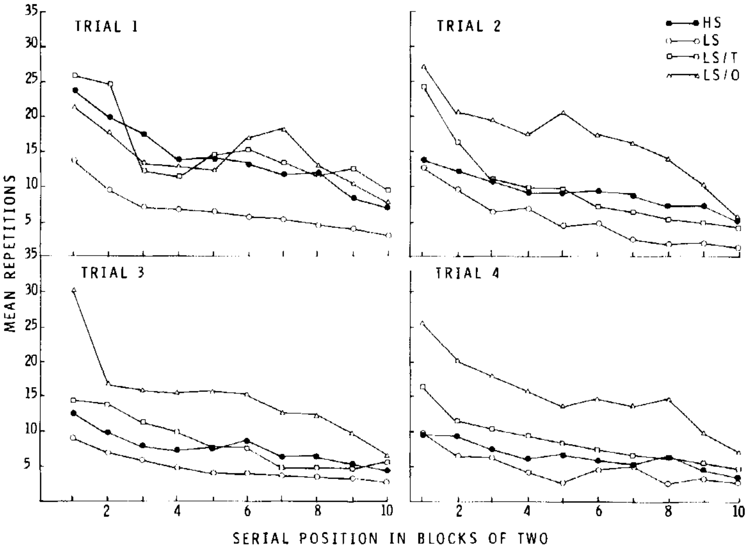


Fig. 16. Repetitions as a function of blocks of two serial positions, for HS and three groups for LS.

quence. The data for the rehearsal-set-size analysis which presents the number of different items rehearsed at each serial position are shown in Fig. 17 where there seems to be a result similar to the others. HS created larger rehearsal sets incorporating more items at each position on Trial 1, then HS gradually became indistinguishable from LS or LS/T, and LS/O continued to create larger rehearsal sets in the longer intervals they were taking to rehearse. In summary, then, LS/O subjects who had been instructed to recall the items in the same order from trial to trial rehearsed for longer periods, produced more repetitions of each item, and a larger number of items in each rehearsal set, at each serial position.

Finally, let us examine the free-recall performance produced by these different rehearsal patterns. The average number of items recalled per trial was 16.2 for HS, 15.6 for LS/O, 14.5 for LS/T, and 14.1 for LS. The trial by trial data for these four groups are shown in Fig. 18. It is generally apparent that LS/O and HS performed approximately the same, as did LS and LS/T for two trials after which LS/T condition improved somewhat. The analysis of variance of these data revealed a significant interaction between Trials and Conditions, $F(12,150) = 3.340$, $p < .01$, as well as a main effect of Trials.

The following points seem worth mentioning concerning the extensive data presented immediately above. First, it appears as though our instructions to LS/O were successful in producing multitrial free recall which was very like that of HS. Nevertheless, there were some basic dif-

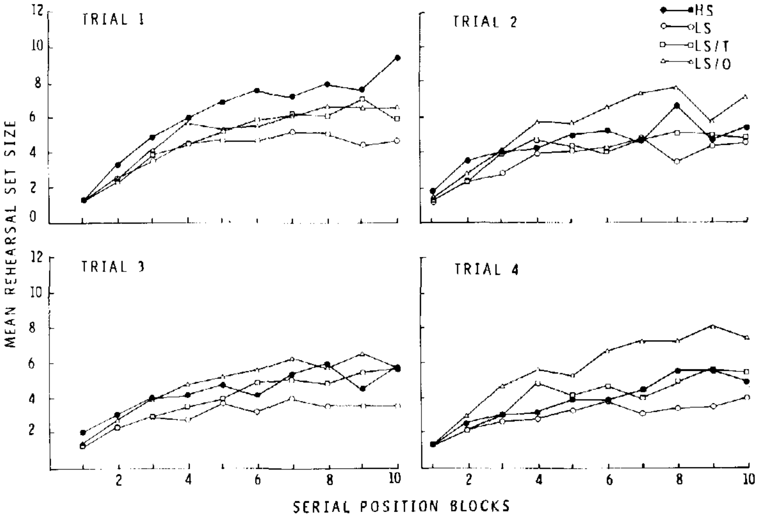


Fig. 17. Mean rehearsal-set size as a function of blocks of two serial positions.

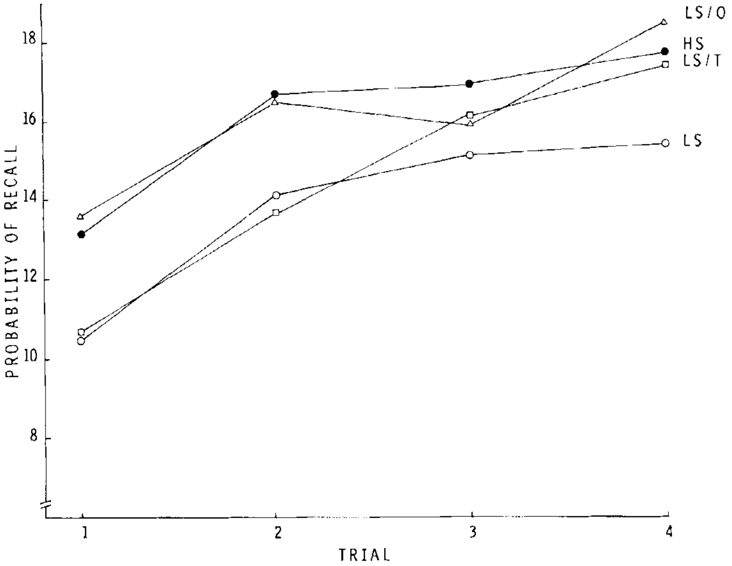


Fig. 18. Free recall as a function of trials for HS and three LS conditions.

ferences. The pattern of behavior for HS might be described something like this. Having rehearsed for long times, including many repetitions, and large rehearsal set sizes once, on Trial 1, HS appeared to have established memory traces of a quality where they could reduce their rehearsal times to the level of LS, reduce the number of repetitions to a minimum, reduce the rehearsal set size substantially, and increase their performance considerably from trial to trial. The pattern of behavior for LS/O might be described as such. Having rehearsed for long times, including many repetitions, and larger rehearsal set sizes on Trial 1, in order to maintain the performance at HS levels, they continued to rehearse for long times, with many repetitions and large set sizes. One intriguing data point is the drop in the performance curve of LS/O on Trial 3 which was preceded by an average rehearsal-set-size drop of 14% on that trial. The point is that although LS can perform as well as HS, given serial instructions, and lots of time to rehearse, there seems to be a cost, and that cost seems to be the maintenance of a much higher level of rehearsal on all trials. In view of these comments, it is difficult to resist the speculative, but somewhat data-based conclusion that HS superiority in this free-recall performance happened as a result of mental events different to those which allowed LS/O to equal them. As tangential evidence as to what this difference might be, we offer Fig. 19, which shows LS and HS performance as a function of the subjective serial-position

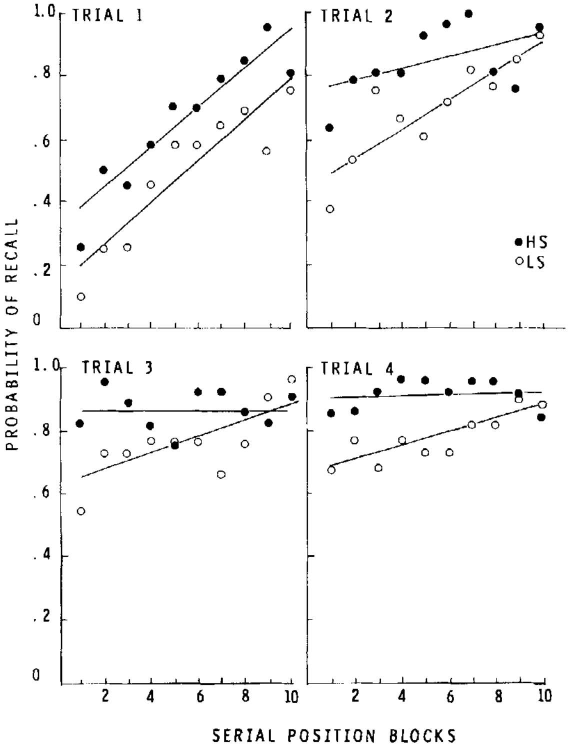


Fig. 19. Subjective serial position curves for LS and HS. Shown are the probabilities of recall as a function of the order in which items were dropped from the rehearsal set.

curve (cf. Brodie & Murdock's 1977 functional serial-position curve). Figure 19 shows the trial-by-trial probability of recall of the 20 items according to the order in which their rehearsal was terminated.[7] The increase in slope on Trial 1 indicates that both LS and HS recall was dependent upon the recency with which the item has been rehearsed overtly. By Trial 3, for HS, recall was no longer based upon recency of rehearsal because the curve is essentially flat, whereas for LS, recency of rehearsal appears to continue to determine the probability of recall. LS

---

[7] Bruce Clarke suggested this analysis spontaneously. It differs from the functional serial position curve in the following way. The functional serial position curve plots the probability of recall of an item against the last serial position in which it was included in the rehearsal set, whereas the subjective serial position curve plots probability of recall of the first item dropped from the rehearsal set as the first point, the probability of recall of the last item said against the last point and so on. In some sense, these data in Fig. 18 depict probability of recall as a function of psychological recency rather than of temporal serial position.

recall remains dependent upon psychological recency longer than HS memory.

IX. When Some People Remember Better Than Others and Perhaps Why

The title of this article promised the reader an account of when some people remember better than others. Since there were 11 separate experiments reported here, and several replications for some, it would be of some benefit, perhaps, to summarize the evidence presented above before offering the arguments about which we think the data are trying to tell us about LS and HS memories. In the original series of four studies, we found that HS remembered better than LS during multitrial free recall, serial recall, and paired-associate learning. LS performance seemed more dependent upon the associations built into the order in which the items were presented, those associations being based upon the preferences, i.e., upon the collective subjective orderings found in the recall protocols of previous free-recall learners. At first, it appeared as though the advantage of the HS was a facility in the formation of interitem associations, but a strong version of this hypothesis has been provisionally abandoned for many reasons. One of those reasons is the finding that PA learning and recognition memory produced similar patterns of differences as free and serial recall; another is the finding that HS could not perform the cued-recall task after Type I rehearsal, a task which required them to provide the second member of a pair of words they had rehearsed together as many as 10 times, given the first member of the pair.

The finding that HS *recognize* better than LS was obtained several times in the present experiments; it occurred after Type I rehearsal, after frequency judgments, after preferred lists and nonpreferred lists, and after processing for case, rhyme, and meaning (both category and sentence judgments) for items presented once and for items presented a number of times.[8] That HS *free recall* better than LS has been illustrated for pre-

---

[8] It is pointed out here, belatedly, that these data add to the weight of evidence on the side of those who argued 10 years ago that organizational variables do influence recognition memory (e.g., Bower, Clark, Lesgold, & Winzenz, 1969; Mandler, 1972). The result that LS consistently managed to *recognize* better after seeing the words in an order which had been preferred for free recall by previous learners seemed remarkable in view of theories of recognition memory which denied any part for organization (subjective or not) in the recognition process. In fact, the mood of that time was so reluctant to publish data which showed "odd" organization effects upon recognition memory that the original submissions based on these data were rejected; the best reason given was that the results might have been a "fluke." Now, 10 years later, it seems self-evident that the present results should have occurred.

ferred and nonpreferred orders of presentation, for free-recall transfer (both part/whole or whole/part learning, Earhard, 1974), for several types of processing, after frequency judgments, and after extended overt rehearsal where LS and HS were given the same task with the same instructions and rehearsed for the same amount of time or more per serial position, for items presented more or less frequently from one to ten times. In addition to the differences found in the tasks mentioned heretofore, one might add alphabetically-cued and categorically-cued recall where significantly fewer items were retrieved by LS than by HS even though it could be demonstrated that many more items were available to LS than were retrieved.

Let us now consider when some people do not remember better than others. There are several important aspects in which LS and HS behave in a similar manner. First, lag controlled the immediate free recall of LS and HS to the same extent over 0 to 32 intervening items. Furthermore, items which subjects reported to have been repeated showed the lag effect, and items which subjects reported to have occurred only once did not, and this was true for both LS and HS. Immediate free recall varied directly with the amount of time for which each item was presented (1.0, 2.5, or 4.0 sec), but this effect was obtained for LS and HS performance. Frequency judgments increased the more separately the repeated instances of a particular word were spaced, and this increase took place at a similar rate for LS and HS. Recall for LS and HS increased at a similar rate with the frequency with which an item had been experienced. The difference between LS and HS was constant for case, rhyme, category, and sentence judgments, indicating that levels of recall after making different levels or different outcomes were *relatively* the same. Furthermore, the relative levels of recall according to alphabetic and category cues are similar, that is LS and HS both remember fewer items when unexpected initial-letter cues are presented than when unexpected category cues are presented. We have some evidence also that the relation between the number of items per cue and the probability of recall is similar for LS and HS. Earhard (1976b, 1974; Ozier, 1977) has reported a very striking relation between alphabetic-cued and category-cued recall and the number of items correctly recalled; the more items there are which begin with a different letter of the alphabet and the more items there are which belong to the same category (up to six or eight items per cue), the less successful is the initial letter or the category name as a retrieval cue. When category-cued and alphabetic-cued recall protocols of LS and HS were scored for the number of items correctly recalled according to the number of items in the list which belonged to the same cue, the difference between LS and HS was a constant. Finally, LS and HS do not differ in retention of CVCs or digit groups after rehearsal which requires the minimum in cognitive capacity or the maximum in cognitive capacity; neither do they differ in the speed of retrieval from semantic memory or in the spread of activation therein.

There have been some times, however, when variables have worked differently for LS and HS. The conditional probability of correctly judging whether a word had been repeated or not was not related to lag for HS but deteriorated for LS over lags from 0 to 32. The lag effect, while present during immediate free recall for both LS and HS, remained clearly present in the delayed free recall of the four lists for HS, but disappeared completely for LS. HS seemed to be able to rehearse well once, using long periods of time, producing many repetitions, and large rehearsal sets, and then to be able to improve their performance without much more activity. LS, when they produced recall levels for HS, did so by continuing to rehearse on subsequent trials, as they did on Trial 1. Finally, an effect we have called psychological recency seemed to have a much longer effect upon LS performance. After one trial HS could drop items from their rehearsal sets without that affecting the probability of recall.

Although this article only promised to tell the reader when some people remember better than others, it is critically important that we try to detect in the numerous findings described above, some guidance as to what might be the locus of the differences in subjective organization reflected in the stereotypy of sequential output found in the recall protocols of our subjects during multitrial free recall. Hunt (1978) has provided an extremely useful organization for our thinking on this matter. Hunt argues that there are three facets of normal cognitive function. The first of these is the knowledge base of the individual based upon past experience. Examples given are logic and mathematics as methods of solving problems. We might add mnemonic strategies to this list. The second aspect of cognitive function Hunt terms mechanistic, and he likens this type to the engineering component of the computer. Hunt distinguishes between mechanistic functions as those which require attention, and those which do not; he instantiates this latter type by the syntactic rules which underly the encoding, decoding, and transmitting of any language. We would add to this a list of mechanistic-automatic memory phenomena including: the spacing/lag effect, the monitoring of frequency of experience, the difference between recall of items given positive and negative judgments, the superior retention of items given semantic processing, primacy and recency effects, the serial position curve, the automatic retrieval from semantic memory of the meanings of things, recognition after Type I processing, the relation between frequency and recall. Memory phenomena which seem to acquire some intention or attention which might fall into this category are recalled according to initial-letter cues, and the function relating probability of recall to the number of items per cue. It is often difficult to decide definitely into which of the two mechanistic categories a memory phenomenon ought properly to be placed. The distribution into these two categories offered here is based upon whether the effect can be avoided, and whether alterations other than increases or decreases in main effects have been commonly reported. Finally, Hunt proposes a third facet to cognitive function, "general information processes, simple strategies that are used as steps in virtually every larger problem" (p. 128). The assumption here is that most problems can be solved in a variety of ways; each individual develops a repertoire of ways of calling into play the routines (e.g., aspects of rehearsal we have examined) that might be effective. Solutions of analogies, styles of playing chess, might be some overt expressions of this third category of influence upon cognitive processes.

Can we, with this organization for our thinking, locate the HS ID we have been studying? It is suggested here that we can dismiss the first category, i.e., differences in basic knowledge, or at least put this aside at the juncture. All of our subjects were normally functioning high school graduates who had manifest above average achievement in order to enter the University. Furthermore, it is suggested that we can dismiss the third category as well. Even when our LS/T subjects were using the elementary rehearsal routines of repeating items over and over, including more items in a set, and spending more time, they did not recall as well as HS for at least two trials. It seems extremely likely that we are dealing with a difference of a mechanistic nature (like other apparently "mechanistic" parameters of the human memory system including: the operation of lag upon recall and frequency judgments, the time course of short-term memory, memory span for digits, retrieval effects of different types of initial processing, retrieval characteristics of traces established during maintenance rehearsal), and extremely likely that it is one which does not require attention since it occurred after Type I rehearsal, to repeat, when there was no intent to memorize, there was no trace for free recall, and there was minimal cognitive capacity devoted to the rehearsal activity.

Any further elaboration of the specifics of the mechanistic ID proposed here is obviously problematic, interpretive, and necessarily conjecture. All that being the case, our best hunch is that HS automatically establish episodic traces which are (for want of a more precise term) of a better *quality*, traces which may be clearer, thus richer, and particularly traces which may be more unique, i.e., *in a context specific to the particular experience being recorded* than do LS. A colleague has named this the "foggy-trace hypothesis" of LS memory function, a phrase appropriately coined on the northeast coast, on the Atlantic Ocean. In the fog, there is only a vague notion of place and time, it is difficult to discriminate one type of ship (indeed one ship) from another, one landmark from another, even though one is generally aware that shadows are passing, and even though one might be able to recognize which landmark is ahead, given several alternative possibilities.

The notion that episodic traces might differ in quality from one individual to another, and that one might be able to study such a variable behaviorally may not be so far fetched as one might imagine at first. Certainly, it is implicit to most of our beliefs about memory that traces of different qualities exist *within* the same individual; otherwise, how would we conceive of successful recognition and successful retrieval for one item, successful recognition and retrieval failure for another item, and recognition failure and retrieval failure for another item, by the same individual, for items which have all been processed to the same semantic level within a very short span of time prior to test (cf. Tulving & Watkins, 1975)? Furthermore, that traces differ in quality *between* individuals (at least between one individual and the rest) has been documented clearly for us by Luria (1968) in his gripping monograph describing the remarkable memory characteristics of the Russian mnemonist S. For S, every experience was unique, and attendant upon each were "synesthetic components [which] were important to his recall, for they created, as it were, a background for each recollection, furnishing him with additional, 'extra' information that would guarantee recall" (p. 28). If the reader will grant, for the sake of argument, that the quality of S's memory traces represents one end of a continuum of uniqueness of experience, related to the accessibility of individual traces, then we might be tempted to advance the argument further and suggest that HS and LS occupy different positions on that continuum.

Relevant to the suggestion that the quality of memory traces has memorial consequences are some data of Sternberg (1970) who examined the effect (upon serial scanning) of degrading the quality of the visual stimulus presented for comparison. Since the quality of the stimulus influenced not only the length of the encoding stage, but also the duration of comparison operations, Sternberg argued that the quality of the stimulus can influence the quality of the *output* of the encoding operations. We have been calling the output of the encoding operations in our experiments the episodic memory trace. And, we see no reason to believe that those outputs could not differ in quality between individuals, even in the absence of actual differences in the physical qualities of the visual stimuli.

Finally, in this regard, we present the position of Friedman (1979) who (after Kolers, 1975) distinguishes between feature detection and feature analysis modes of perceptual analysis to differentiate between automatic registration of experience that is anticipated, on the one hand, and more active detailed analysis of less usual experience requiring greater attention for comprehension. We believe that these different modes of perceptual analysis would be expected to result in different qualities of memorial residue, and Friedman's data certainly suggest that feature analysis results in better retention.

This article began with the phenomenon of subjective organization, the finding that, in multitrial free recall, the degree to which items are recalled in the same order from trial to trial is related directly to successful remembering. What is being suggested here is that as an outcome of individual differences in the quality of the episodic memory trace, there might occur a greater success in rearranging the traces mentally, i.e., in manipulating the material to be recalled independent of the order in which the traces were originally experienced, and consonant with other arrangements that might improve the probability of recall. For example, we repeat the notion that clearer, more unique traces would be more likely to be arranged in implicit mental contiguity with one another, according to some attribute that the two mental representations had in common, or cross-classified with one another—so that the order we see, in the output of HS, reflects indirectly the greater richness of the original trace. Foggier traces would deteriorate naturally more quickly, and, in the absence of integration, they would require continued rehearsal in order to achieve the same level of performance by a different routine. One might even argue that foggier traces would seem more similar to one another, and might even be mistaken for one another, thus the higher estimates by LS of the frequency with which the repeated items had occurred. This discussion has veered dangerously far from our data, and it may by now seem somewhat remote from the findings of our experiments. What we are suggesting is that our results will fit the tale told. There may be other stories just as fitting; the tests of this one remain to be made.

REFERENCES

Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review*, 1972, **79**, 97-123.

Anderson, J. R., & Bower, G. H. A propositional theory of recognition memory. *Memory and Cognition*, 1974, **2**, 406-412.

Baddeley, Alan D. The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review*, 1978, **85**, 139-152.

Battig, W. F. The flexibility of human memory. In L. S. Cermark & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, N.J.: Erlbaum, 1979. Pp. 23-44.

Begg, I. Estimation of word frequency in continuous and discrete tasks. *Journal of Experimental Psychology*, 1974, **102**, 1046-1052.

Bellezza, F. S., Richard, D. L., & Geiselman, R. E. Semantic processing and organization in free recall. *Memory and Cognition*, 1976, **4**, 415-421.

Bevan, W., Dukes, W., & Avant, L. The effect of variation in specific stimuli on memory for their superordinates. *American Journal of Psychology*, 1966, **79**, 250-257.

Bower, G. H. Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding process in human memory*. Washington, D.C.: Winston, 1972.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 323-343.

Bower, G. H., Lesgold, A. M., & Tieman, D. Grouping operations in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 501-506.

Bransford, J. D., Franks, J. J., Morris, C. D., & Stein, B. S. Some constraints on learning and memory research. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, N.J.: Erlbaum. 1977.

Brodie, D. A., & Murdock, B. B., Jr. Effect of presentation time on nominal and functional serial position curves of free-recall. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 185-200.

Brodie, D. A., & Prytulak, L. S. Recall curves: Nothing but rehearsing some items more or recalling them sooner? *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 549-564.

Buschke, H. Learning is organized by chunking. *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 313-324.

Buschke, H. Two-dimensional recall: Immediate identification of clusters in episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 201-215.

Cermak, L. S., & Craik, F. I. M. *Levels of processing in human memory*. Hillsdale, N. J.: Erlbaum, 1979.

Chiang, A., & Atkinson, R. C. Individual differences and interrelationships among a set of cognitive skills. *Memory and Cognition*, 1976, **4**, 661-673.

Cofer, C. N. *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961.

Collins, A. M., & Loftus, E. F. A spreading-activation theory of semantic processing. *Psychological Review*, 1975, **82**, 407-429.

Coltheart, V. Recognition errors after incidental learning as a function of different levels of processing. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, **3**, 437-444.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 671-684.

Craik, F. I. M., & Tulving, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, **104**, 268-294.

Cuvo, A. J. Developmental differences in rehearsal and free recall. *Journal of Experimental Child Psychology*, 1975, **19**, 265-278.

Dean, J., & Ley, R. Effects of associative encoding on free recall in high and low verbal associators. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, **3**, 316-324.

Earhard, M. Subjective organization and list organization as determinants of free recall and serial recall memorization. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 501-507.(a)

Earhard, M. Cued recall and free recall as a function of the number of items per cue. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 257-263.(b)

Earhard, M. Individual differences in subjective organization: Short-term memory. *Psychonomic Science*, 1970, **18**, 220-221.(a)

Earhard, M. *Recognition memory, order of presentation, and subjective organization.* Paper presented at the Annual Meeting of the Eastern Psychological Association, April 1970.(b)

Earhard, M. Category cued recall and free recall as a function of the number of items per cue. *Canadian Journal of Psychology,* 1972, **26,** 11–18.

Earhard, M. Individual differences in free recall transfer. *Journal of Experimental Psychology,* 1974, **103,** 1169–1174.

Earhard, M., & Endicott, O. R. Why are there individual differences in subjective sequential organization during free recall memorization? *Journal of Verbal Learning and Verbal Behavior,* 1969, **7,** 316–319.

Ebbinghaus, H. *Über das Gedächtnis.* Leipzig: Duncker & Humblot, 1885.

Eysenck, M. *Human memory: Theory, research and individual differences.* New York: Pergamon, 1977.

Eysenck, M. W. Levels of processing: A reply to Lockhart and Craik. *British Journal of Psychology,* 1978, **69,** 177–178.

Flexser, A. J., & Bower, G. H. Further evidence regarding instructional effects on frequency judgments. *The Psychonomic Society,* 1975, **3,** 321–324.

Folkard, S. Time of day and level of processing. *Memory and Cognition,* 1979, **7,** 247–252.

Freedman, J. L., & Loftus, E. F. Retrieval of words from long-term memory. *Journal of Verbal Learning and Verbal Behavior,* 1971, **10,** 107–115.

Friedman, A. Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General,* 1979, **108,** 316–355.

Friendly, M. L. In search of the M-gram: The structure and organization in free recall. *Cognitive Psychology,* 1977, **9,** 188–249.

Gagne, R. M. (ed.). *Learning and individual differences.* Columbus: Merrill, 1967.

Gagnon, M., Cavanagh, P., & Laurencelle, L. Processing differences between memory search and foveal visual search. *Perception and Psychophysics,* 1978, **23,** 258–260.

Glenberg, A. M. Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, **3,** 282–292.

Glenberg, A., & Adams, F. Type I rehearsal and recognition. *Journal of Verbal Learning and Verbal Behavior,* 1978, **17,** 455–463.

Glenberg, A., Smith, S. M., & Green, C. Type I rehearsal: Maintenance and more. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 339–352.

Goldman, S. R., & Pellegrino, J. W. Processing domain, encoding elaboration, and memory trace strength. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 29–44.

Gorfein, D., Blair, C., & Rowland, C. The generality of free recall: I. Subjective organization as an ability factor. *Psychonomic Science,* 1968, **11,** 279–280.

Hasher, L., & Chromiak, W. The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 173–184.

Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.

Hintzman, D. L. Apparent frequency as a function of frequency and the spacing of repetitions. *Journal of Experimental Psychology,* 1969, **80,** 139–149.

Hintzman, D. L. Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium.* Potomac, Md.: Erlbaum, 1974.

Hintzman, D. L. Repetition and memory. In G. Bower (Ed.), *The psychology of learning and motivation.* New York: Academic Press, 1976.

Hintzman, D. L., & Block, R. A. Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology,* 1971, **88,** 297–306.

Hintzman, D. L., & Bock, R. A. Memory for the spacing of repetition. *Journal of Experimental Psychology,* 1973, **99,** 70–74.

Hintzman, D. L., & Stern, L. D. Contextual variability and memory for frequency. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, **4,** 539-549.

Hopf-Weichel, R. Reorganization in semantic memory: An interpretation of the facilitation effect. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 261-275.

Horton, K. D. Phonemic similarity, overt rehearsal, and short-term store. *Journal of Experimental Psychology: Human Learning and Memory,* 1976, **3,** 244-251.

Howell, W. C. Storage of events and event frequencies: A comparison of two paradigms in memory. *Journal of Experimental Psychology,* 1973, **98,** 260-263.(a)

Howell, W. C. Representation of frequency in memory. *Psychological Bulletin,* 1973, **80,** 44-53.(b)

Hultsch, D. F. Organization and memory in adulthood. *Human Development,* 1971, **14,** 16-29.

Hunt, E. B. Quote the raven? Nevermore! In L. W. Gregg (Ed.), *Knowledge and cognition.* Hillsdale, N.J.: Erlbaum, 1974. Pp. 129-157.

Hunt, E. Mechanics of verbal ability. *Psychological Review,* 1978, **85,** 109-130.

Hunt, E., Frost, N., & Lunneborg, C. Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 7). New York: Academic Press, 1973.

Hunt, E., Lunneborg, C., & Lewis, J. What does it mean to be high verbal? *Cognitive Psychology,* 1975, **7,** 194-227.

Hyde, T. S., & Jenkins, J. J. Recall for words as a function of semantic, graphic, and syntactic orientating tasks. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 471-480.

Jablonski, E. M. Free recall of children. *Psychological Bulletin,* 1974, **81,** 522-539.

Jacoby, L. L. The role of mental contiguity in memory: Registration and retrieval effects. *Journal of Verbal Learning and Verbal Behavior,* 1974, **13,** 483-496.

Jacoby, L. L., Barty, W. H., & Evans, J. D. A functional approach to levels of processing. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, **4,** 331-346.

James, William *Psychology.* New York: Holt, 1892.

Jenkins, J. J. Comments on Prof. Noble's paper. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior.* New York: McGraw-Hill, 1961. Pp. 146-150.

Jenkins, J. J. Individual differences in verbal learning. In R. M. Gagne (Ed.), *Learning and individual differences.* Columbus, Ohio: Merrill, 1967. Pp. 45-57.

Jenkins, J. J. Remember that old theory of memory? Well, forget it! In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing.* Hillsdale, N.J.: Erlbaum, 1977. Pp. 413-429.

Jensen, A. R. Individual differences in visual and auditory memory. *Journal of Educational Psychology,* 1971, **62,** 123-131.

Jusczyk, P. W. Rhymes and reasons: Some aspects of the child's appreciation of poetic form. *Developmental Psychology,* 1977, **13,** 599-607.

Keller, D., & Kellas, G. Typicality as a dimension of encoding. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, **4,** 78-85.

Kintsch, W. Recognition and free recall of organized lists. *Journal of Experimental Psychology,* 1968, **78,** 481-487.

Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of Human Memory.* New York: Academic Press, 1970.

Kolers, P. Memorial consequences of automatized encoding. *Journal of Experimental Psychology: Human Learning and Memory,* 1975, **1,** 689-701.

Kolers, P. A pattern-analyzing basis of recognition. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, N.J.: Erlbaum, 1979. Pp. 363-384.

Landauer, T. K. Memory without organization: Properties of a model with random storage and undirected retrieval. *Cognitive Psychology,* 1975, **7,** 495-531.

Ley, R., & Dean, J. Effects of subject-generated and experimenter-supplied associations at time of storage on free recall of high and low verbal associators. *Psychological Reports,* 1976, **38,** 499-506.

Locascio, D., & Ley, R. Associative reaction time and meaningfulness in free recall. *Journal of Experimental Psychology,* 1972, **95,** 460-462.

Lockhart, R. S. Remembering events: Discussion of papers by Jacoby & Craik, Battig & Nelson. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, N.J.: Erlbaum, 1979. Pp. 77-85.

Lockhart, R. S., Craik, F. I. M., & Jacoby, L. L. Depth of processing, recognition and recall: Some aspects of a general memory system. In J. Brown (Ed.), *Recall and recognition.* London: Wiley, 1976.

Loftus, E. F. Category dominance, instance dominance, and categorization time. *Journal of Experimental Psychology,* 1973, **97,** 70-74.

Loftus, E. G., & Cole, W. Retrieving attribute and name information from semantic memory. *Journal of Experimental Psychology,* 1974, **102,** 1116-1122.

Loftus, G. R., & Loftus, E. F. The influence of one memory retrieval on a subsequent memory retrieval. *Memory and Cognition,* 1974, **2,** 467-471.

Luria, A. R. *The mind of a mnemonist.* New York: Avon, 1968.

Lyon, D. R. Individual differences in immediate serial recall: A matter of mnemonics? *Cognitive Psychology,* 1977, **9,** 403-411.

McGovern, J. G. Extinction of associations in four transfer paradigms. *Psychological Monographs,* 1964, **78.**

McLeod, C. M., Hunt, E., & Mathews, N. Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior,* 1978, **17,** 493-507.

Madigan, S. A. Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior,* 1969, **8,** 828-835.

Mandler, G. Organization and recognition. In E. Tulving and W. Donaldson (Eds.), *Organization of memory.* Academic Press, New York, 1972. Pp. 139-166.

Martin, E. Stimulus meaningfulness and paired-associate transfer: A encoding-variability hypothesis. *Psychological Review,* 1968, **75,** 421-441.

Mathews, R. C., & Tulving, E. Effects of three types of repetition on cued and non-cued recall of words. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 707-721.

Melton, A. W. Repetition and retrieval from memory. *Science,* 1967, **158,** 532.

Melton, A. W. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 596-606.

Miller, G. A. *The subjective lexicon.* Paper presented at the 78th Annual Meeting of the American Psychological Association, Miami Beach, September 1970.

Murdock, B., & Metcalfe, J. Controlled rehearsal in single trial free recall. *Journal of Verbal Learning and Verbal Behavior,* 1978, **17,** 309-324.

Nelson, T. O. Reception and depth of processing. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 151-171.

Ornstein, P., Naus, M. J., & Liberty, C. Rehearsal and organizational processes in chidlren's memory. *Child Development,* 1975, **46,** 818-830.

Ozier, M. Retrieval failure in the presence of retrieval cues: A comparison of three age groups. *Canadian Journal of Psychology,* 1977, **31,** 139-150.

Ozier, M. Access to the memory trace through orthographic and categoric information. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, **4,** 469-485.

Ozier, M. *Access to names and meaning of words in semantic memory by preteens, young adults, and senior adults.* Paper presented at the Annual Meeting of the Eastern Psychological Association, Philadelphia, April 1979.

Ozier, M., & Anderson, B. *Individual differences in free recall: A memory trace strength hypothesis.* Paper presented at the Annual Meeting of the American Psychological Association, New York, August 1979.

Ozier, M., & Sperry, J. *Individual differences in subjective organization.* Paper presented at the Annual Meeting of the Canadian Psychological Association, Ottawa, June 1978.

Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology,* 1959, **58,** 193-198.

Plenderleith, Marus, & Postman, L. Individual differences in intentional and incidental learning. *British Journal of Psychology,* 1957, **48,** 241-248.

Postman, L. In C. N. Cofer, *Verbal learning and verbal behavior.* New York: McGraw-Hill, 1961. P. 151.

Postman, L., & Kruesi, E. The influence of orienting tasks on the encoding and recall of words. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 353-369.

Rose, R. J., & Rowe, E. J. Effects of orienting task and spacing of repetitions on frequency judgments. *Journal of Experimental Psychology: Human Learning and Memory,* 1976, **2,** 142-152.

Rowe, E. J. Frequency judgments and recognition of homonyms. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 440-447.

Rowe, E. J. Depth of processing in a frequency judgment task. *Journal of Verbal Learning and Verbal Behavior,* 1974, **13,** 638-643.

Rowe, E. J., & Rose, R. J. Effects of orienting task, spacing of repetitions, and list content on judgments of frequency. *Memory and Cognition,* 1977, **5,** 505-512.

Rozin, P., & Gleitman, L. The structure and acquisition of reading II: The reading process and the acquisition of the alphabetic principle. In A. S. Reber & D. Scarborough (Eds.), *Reading: The C.U.N.Y. Conference.* Potomac, Md.: Erlbaum, 1977.

Rundus, D. Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology,* 1971, **89,** 63-67.

Rundus, D. Maintenance rehearsal and single-level processing. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16,** 665-681.

Rundus, D., & Atkinson, R. C. Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 99-105.

Runquist, W. N., & Runquist, P. A. Interference reduction with conceptually similar paired associates. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, **4,** 370-381.

Sakoda, J. M. Individual differences in correlation between clustering and recall of meaningful words. *Journal of General Psychology,* 1956, **54,** 183-190.

Schulman, A. I. Memory for words recently classified. *Memory and Cognition,* 1974, **2,** 47-52.

Shapiro, S. I., & Bell, J. A. Subjective organization and free recall: Performance of high, moderate, and low organizers. *Psychonomic Science,* 1970, **21,** 71-72.

Shipley, W. C. A self-administering scale for measuring intellectual impairment and deterioration. *The Journal of Psychology,* 1940, **9,** 371-377.

Sternberg, R. J., & Tulving, E. The measurement of subjective organization in free recall. *Psychological Bulletin,* 1977, **84,** 539-556.

Sternberg, S. Memory scanning: Mental processes revealed by reaction time (RT) experiments. In J. S. Antrobus (Ed.), *Cognition and affect.* Boston: Little, Brown, 1970.

Thorndike, E. L., & Lodge, I. *The teacher's word book of 30,000 words.* New York: Columbia University, Teachers College, Bureau of Publications, 1944.

Tulving, E. Subjective organization in free recall of "unrelated" words. *Psychological Review,* 1962, **69,** 344-354.

Tulving, E. Ecphoric processes in recall and recognition. In J. Brown (Ed.), *Recall and recognition.* New York: Wiley, 1976. Pp. 37-73.

Tulving, E., & Watkins, M. S. Structure of memory traces. *Psychological Review,* 1975, **82,** 261-275.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, **5**, 207-232.

Underwood, B. J. Some correlates of item repetitions in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 83-94.

Underwood, B. J. Word recognition memory and frequency information. *Journal of Experimental Psychology*, 1972, **94**, 276-283.

Underwood, B. J. Individual differences as a crucible in theory construction. *American Psychologist*, 1975, **30**, 128-134.

Underwood, B. J., & Malmi, R. A. An evaluation of measures used in studying temporal codes for words within a list. *Journal of Verbal Learning and Verbal Behavior*, 1978, **3**, 279-293.

Woodrow, H. The ability to learn. *Psychological Review*, 1946, **53**, 147-158.

Woodward, A. E., Bjork, R. A., & Jongeward, R. H. Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 608-617.

INDEX