HANDBOOK OF LEARNING AND COGNITIVE PROCESSES

6 Volume set

Edited by
W. K. Estes

Psychology Press

INTRODUCTION TO CONCEPTS AND ISSUES

Volume 1

Edited by
W. K. Estes

HANDBOOK OF LEARNING AND COGNITIVE PROCESSES

Psychology Press

HANDBOOK OF LEARNING AND COGNITIVE PROCESSES

Volume 1

INTRODUCTION TO CONCEPTS AND ISSUES

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

HANDBOOK OF LEARNING AND COGNITIVE PROCESSES

Volume 1

Introduction to Concepts and Issues

Edited by
W. K. ESTES

Psychology Press
Taylor & Francis Group
LONDON AND NEW YORK

First published in 1975

This edition first published in 2014
by Psychology Press
27 Church Road, Hove BN3 2FA

and by Psychology Press
711 Third Avenue, New York, NY 10017

Psychology Press is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-84872-347-4 (Set)
eISBN: 978-1-315-76974-5 (Set)
ISBN: 978-1-84872-388-7 (Volume 1)
eISBN: 978-1-315-78107-5 (Volume 1)

Publisher's Note
The publisher has gone to great lengths to ensure the quality of this book but points out that some imperfections from the original may be apparent.

Disclaimer
The publisher has made every effort to trace copyright holders and would welcome correspondence from those they have been unable to trace.

Handbook of Learning and Cognitive Processes

Volume 1

Introduction to Concepts and Issues

EDITED BY

W. K. ESTES

Rockefeller University

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1975 Hillsdale, New Jersey

DISTRIBUTED BY THE HALSTED PRESS DIVISION OF

JOHN WILEY & SONS

New York Toronto London Sydney

Lawrence Erlbaum Associates, Inc., Publishers
62 Maria Drive
Hillsdale, New Jersey 07642

Distributed solely by Halsted Press Division
John Wiley & Sons, Inc., New York

Library of Congress Cataloging in Publication Data
Main entry under title:

Handbook of learning and cognitive processes.

Includes bibliographical references and indexes.
CONTENTS: v. 1. Introduction to concepts and issues.
1. Cognition 2. Learning, Psychology of. I
I. Estes, William Kaye. [DNLM: 1. Cognition.
2. Learning. LB1051 H236]
BF311.H334 153 75-20113
ISBN 0-470-24585-9

Printed in the United States of America

Contents

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Foreword

Is it possible at present to identify a core cluster of theoretical ideas, concepts, and methods with which everyone working in the area of learning and cognition needs to be familiar? Would it be possible to make explicit the relationships that we feel do or must exist among the various subspecialties, ranging from conditioning through perceptual learning and memory to psycholinguistics, and to present these in a sufficiently organized way to help specialists and nonspecialists alike in relating particular lines of research to the broader spectrum of activity?

These questions were posed to a substantial number of the investigators who are currently most active in developing the ideas and doing the research. Their response constitutes this *Handbook*.

Beyond trying to make clear how they think and what they do, the contributors have been invited to address the question of what they hope to accomplish. The problem of long-term objectives is one of increasing concern to many people in the field. Perhaps all investigators believe, or at least hope, that the large and ever-expanding volume of research bears the potentiality of major contributions to the solution of human problems. But a scan of the research literature yields little evidence that the various studies recorded have direct bearing on problems arising outside the laboratory. A reader tends to gain the impression that researchers are primarily oriented toward the development of theoretical accounts of experiments or at most groups of experiments. But problems belong to individuals, not to experiments. Who is thinking about the matter of putting the results together to produce theories that may elucidate the functioning of individuals in the various kinds of situations that call for cognitive or intellectual activities?

The answer is perhaps that, although most investigators have long range goals in mind, there is no place for expression of these in research reports or even in most theoretical articles. Consequently, one of our major purposes in the present undertaking has been to offer a number of experienced investigators an opportunity to present in relatively broad perspective the research areas with which they are closely familiar and to consider explicitly the possibilities of bringing research results to bear both upon the development of more general theories and upon approaches to practical problems.

Another major problem of communication arises from the fact that the volume of research results overflows available communication channels. Therefore, conciseness of report is a major requirement, and as a consequence most of the research literature is accessible only to those with special interest and competence in a specific area. To a considerable extent this state of affairs is only a natural accompaniment of the rapid acceleration in research and the consequent tendencies toward fractionation of the field (Estes, 1972).[1] Research and theory on conditioning and animal learning have been revitalized by the results of electrophysiological studies of the neural systems involved in reinforcement and motivation and to some extent also by developments in ethology. Consequently, investigators of conditioning find it necessary to keep up with technical literature in these areas and to communicate the results primarily to others who are on the same footing. Nonetheless, it is possible that interpretations of conditioning could benefit also from ideas arising in contemporary studies of memory and information processing, and in fact there are some specific indications that such is the case (as witness a number of chapters in Glaser, 1971; Mackintosh & Honig, 1969; and McGuigan & Lumsden, 1973).[2] Similarly, much current work in cognitive psychology has been catalyzed by inputs from linguistics and computer science, again calling for new expertise on the part of those who would keep up with the literature.

The result has been to make current literature in cognitive psychology appear increasingly remote from that of conditioning and learning theory. Nonetheless, however we may choose to view it, cognitive psychology has a biological basis, and there must be serious hazards in isolating the stream of development of cognitive theories from contemporary developments in conditioning and psychobiology.

[1] Estes, W. K. Learning. In P. C. Dodwell (Ed.), *New horizons in psychology 2.* London: Penguin Books, 1972. Pp. 15–35.

[2] Glaser, R. (Ed.), *The nature of reinforcement.* New York: Academic Press, 1971.

Mackintosh, N. J., & Honig, W. K. *Fundamental issues in associative learning.* Halifax, Nova Scotia: Dalhousie University Press, 1969.

McGuigan F. J., & Lumsden, D. B. (Eds.), *Contemporary approaches to conditioning and learning.* Washington, D.C.: Winston, 1973.

The contributors to this *Handbook* have not been satisfied to bewail the thinning of communication among various facets of learning and cognitive psychology, but rather have joined in an effort to do something about the situation. We have attempted to organize and present a picture of the current state of the field that will be fairly up to date with regard to theoretical and even technical developments and yet readable for anyone with a reasonable scientific background, regardless of his acquaintance with the technical jargon of particular specialties.

Working within this one constraint, our first emphasis has been to present the major concepts, theories, and methods with which one should be familiar in order to understand or to participate in research in any of the various facets of cognitive psychology. Second, each of the authors has taken on the assignment of giving explicit attention to the orienting attitudes and long-term goals that tend to shape the overall course of research in his field and to bring out both actual and potential influences and implications with respect to other aspects of the discipline.

To have a chance at realizing our main objectives, we, of course, have had to be severely selective with regard to coverage of factual material. We have made a real effort toward completeness in the sense of giving representation to all of the principal types of research results that have major bearing on theoretical issues, of providing critical discussion on the methodological problems that have had to be solved in order to produce this relevance, and of providing an adequate starting point for those who wish to gain a foothold on the technical literature.

We attempt in Volume 1 of this *Handbook* to present an overview of the field and to introduce the principal theoretical and methodological issues that will persistently recur in the expanded treatments of specific research areas that comprise the later volumes. Deferring to the current Zeitgeist rather than to chronology, we begin with the present state of cognitive psychology, then introduce the comparative approach, and conclude this volume with a rapid, three-chapter review of the evolution of ideas from conditioning to information processing.

The next two volumes will treat the areas traditionally associated with learning theory: Volume 2, conditioning and behavior theory; Volume 3, human learning and motivation. The last three volumes range over the many presently active lines of research identified with human cognitive processes: Volume 4, attention, memory storage and retrieval; Volume 5, organization in memory, cognitive processes in reading; Volume 6, information processing, problem solving, and artificial intelligence. The oganization by volumes is of necessity somewhat arbitrary, but as far as possible the lines of demarcation have been drawn with respect to theory rather than to method. Thus, for example, developmental approaches are represented by chapters in several volumes rather than being collected as a unit.

The completion of a publishing enterprise that initially seemed almost too ambitious to contemplate has required a rather massive cooperative effort. Most apparent are the efforts on the part of a large number of busy and often overcommitted scientists to present their ideas in a way that might cut across the usual boundaries. Bringing these efforts to fruition has involved unusually close and continuing collaboration between those doing the writing and editing and the publisher, Lawrence Erlbaum, over a period beginning many years before the first of these chapters was written. Finally, I am indebted to a number of colleagues, especially Michael Cole, Katherine W. Estes, and Douglas L. Medin, for critical readings of manuscripts and to my assistant, Edith Skaar, for her efforts in helping steer this project through a sea of detail under which the Editor otherwise would long ago have vanished from view.

W. K. ESTES

Handbook of Learning and Cognitive Processes

Volume 1

Introduction to
Concepts and Issues

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

1

The State of the Field: General Problems and Issues of Theory and Metatheory

W. K. Estes

Rockefeller University

I. A PROVISIONAL DEFINITION

Just as the physical sciences can be conceived as the study of energy in its many aspects, the behavioral and social sciences can be characterized in terms of their concern with the processing and transformation of information. The adaptation of living organisms to an ever-changing environment depends upon the ability to acquire information about environmental regularities and to use this information as the basis of adaptive response. This process occurs on two very different time scales. Over the long time scale, of the order of centuries, species acquire and store information in a genetic code—the basis for so-called instinctive or species-specific behavior. Over the short time scale, information is acquired by individuals and stored in the nervous system in a manner still largely unknown. Further, the storage of information by individuals in a capacious but impermanent memory is augmented by various types of records that require the individual to remember only the code and not the contents in order to reproduce indefinitely large quantities of stored information. The scientific study of the processing of information on a short time scale defines the field of learning and cognitive psychology.

The two most conspicuous aspects of cognitive activity are perhaps power and fragility. Successful operation of cognitive functions has yielded civilization as we know it—science, philosophy, mathematics, the arts, commerce, government. But failures are fully as striking. Millions of individuals

never acquire the rudimentary intellectual skills needed for everyday functioning in a social environment. Many more are hampered in their social adjustments by inability to communicate effectively in either speech or writing. Many more are able to acquire intellectual and communicative skills but are frustrated by an inability to comprehend enough of complex social institutions to have a fair chance to share in the benefits. Still others reach relatively high levels of mental functioning only to find their powers slipping away under the influence of disease, drugs, stress, or aging. It is only natural then that an increasing number of thoughtful and socially concerned individuals should be coming to believe that one of the most urgent tasks for our society is the development of a body of scientific knowledge that will enable us to understand more of the capacities and limits of human information processing and to guide efforts to remedy deficiencies or failures of cognitive functioning.

The various subdisciplines engaged in this task may be grouped under the currently popular generic label cognitive psychology. Theoretically, cognitive psychology might be defined in terms of the study of all aspects of information processing by organisms. But to be practical, at least for purposes of the *Handbook,* we need to apply some limitations. Without trying to draw too fine a line, we shall deal primarily with studies which focus on the storage, retrieval, and utilization of information rather than on the neurophysiological mechanisms by which these functions are accomplished. Further, although observed behavior of organisms must be our principal index concerning states of information, we shall largely exclude from our consideration a large body of work of a more or less engineering character in which the methods of operant conditioning, behavior modification, or the motor skills laboratory are utilized for the shaping of behavior as an end in itself.

II. VARIETIES OF RESEARCH

What the study of learning and cognitive processes appears to be depends on where one looks. To illustrate, suppose that we were to inspect on a typical day some of the studies currently being conducted by members of my laboratory group. In one experimental chamber we might see human subjects observing displays of letters on a screen while their eye movements are monitored electrically by an apparatus so contrived that the display disappears when the individual moves his eyes—thus permitting the experimenter to determine how much a person can see in a single look. In another room adult subjects are participating in an experiment in which the apparatus is controlled by a computer so as to simulate the operation of a public opinion poll. A bit further away a study is in progress in which the short-term memory of monkeys is being tested, the investigator hoping to discover whether these animals are able to maintain retention by a process similar to rehearsal. If we proceed still further from the heart of the labora-

tory, we might see a group of bilingual children in a city preschool "playing a game" with an experimenter who seeks to understand how the children acquire the ability to deal with number concepts. And at the same time, though we could not conveniently view the operation, halfway around the globe some of our colleagues are studying the acquisition of similar skills by African children who grow up in a quite different cultural setting.

Some of the studies involve highly sophisticated instrumentation and deal with precise measurement of narrowly circumscribed behaviors that seem remote from anything occurring outside of the laboratory. Others deal with tasks similar to or even drawn directly from practical problems arising outside the laboratory. In some studies the experimenter goes to extreme lengths to prevent the subjects from introspecting. In others, introspection is encouraged, and the reports are recorded as an interesting part of the data.

This wide variation in content, methods, and orientation of researches is typical of a field in which techniques must be found to study phenomena too complex to yield to any one approach and in which the research is directed toward objectives no one yet knows how to realize.

Viewed from the outside, perhaps the most salient features of present-day cognitive psychology are, on the one hand, the volume and heterogeneity of research, and, on the other, the absence of any semblance of plan or order. Trying to read through any year's volumes of research articles is much like walking through a field which has been sown by mixing all varieties of seed, from asters to zinnias and alfalfa to wheat, and scattering them at random. Innumerable small plots are cultivated industriously but autonomously, today's harvest in any one yielding the germ of the next, except perhaps for haphazard wind-blown inputs from neighbors.

This appearance in a way is deceptive. Behind the noisy surface of research output there is actually a substantial measure of plan and order. The order exists, however, not in the activity nor in the products themselves, but only in the mind of the scientifically trained observer who possesses the key to the system. The key is theory. The organizing influence of ideas, concepts, and attitudes generates significant categorizations of phenomena, imposes gradients of subordination, makes it possible to retrieve from diverse sources the information required to solve a problem. Thus, to understand cognitive psychology, one must understand its theories.

III. METATHEORETICAL ISSUES AND ORIENTATIONS

The form taken by a theory is dictated partly by facts, partly by philosophical and methodological presuppositions that constrain and shape the thinking of those who do the theorizing. It is scarcely surprising, therefore, to find that research on learning and cognition has tended to be polarized with respect to a number of issues which are familiar from the histories

of other sciences. One of these is the problem of objectivity, which in psychology is most conspicuously manifest in the persisting opposition of behavioral and cognitive approaches. A second reflects differing views concerning the proper level of generality for psychological laws. A third is the increasingly acute problem of relevance to practical affairs. In this section I shall summarize some ideas that may help relate these issues to the treatments of theory and method that constitute the body of the *Handbook*.

A. Mind and Behavior

Long before experimental psychology made its earliest appearance, other biological and physical sciences had accumulated imposing records of accomplishment with regard to prediction and control of natural phenomena. These achievements were attributable in great measure to the successes of the sciences in achieving objectivity, communicability, and verifiability of findings and reducibility or interconnectedness of their theories. It was natural that any other disciplines that hope for the same successes should strive to emulate the methods.

In study of learning and thought, the problem of objectivity surfaced in the guise of the now nearly century old opposition of introspectionist versus behavioristic approaches. It appeared to the early structuralists (for example Titchener, 1909) that introspections could be as repeatable and measurable phenomena as overt behaviors, provided only that they are obtained under sufficiently rigorously controlled conditions. Repeatability of observations did indeed appear to be achievable, but no method was forthcoming to meet the similarly important requirement of agreement among independent observers; and, perhaps worse, the results of introspective studies proved refractory to any theoretical connection with results of researches in other disciplines.

The behavioristic approach (Watson, 1924), in contrast, seemed on sound ground with respect to these desiderata. Observations and measurements of overt behavior could be made as objective as observations and measurements on pendulums or raindrops and, thanks to Pavlov (1927), the concepts emerging from the observations appeared reducible or at least relatable to those of physiological and neural sciences. In consequence, over the greater part of the period during which the psychology of learning and cognition has taken shape, most investigators have, with different degrees of willingness and enthusiasm, subjected themselves to the discipline of methodological behaviorism.

Contemporaneous with the hegemony of associationism and behaviorism, there was always present in the background an alternative, phenomenologically oriented approach in terms of concepts of mental functioning

drawn from the vocabulary of everyday life. However, harnessed to the introspectionist methodology, conceptions of learning and cognition couched in terms of mental processes did not begin to grow to the stature of formal theories until the recent relaxation of the hold of behavioristic thinking.

Only in the very last few years have we seen a major release from inhibition and the appearance in the experimental literature on a large scale of studies reporting the introspections of subjects undergoing memory searches, manipulations of images, and the like. This disinhibition appears to be a consequence of a combination of factors. Among these are new developments in methodology that help with the problem of verifiability and a broadening body of theory that helps to solve the problem of connectedness with other disciplines. Further, new sources of conceptual analyses, especially those arising from information-processing systems, make mental operations seem less subjective to the extent that they appear more and more to resemble processes that are realizable in machines, or at least in programs that can be run on machines.

The construction of theory in terms of mental processes seems finally to have come into respectability with the advent of digital computers and demonstrations that, by simulating covert processes in computer models, one can achieve both the appearance of greater objectivity and the fact of greater testability. Thus, the present-day investigator of learning and cognition tends to choose between the behavioral and cognitive approaches, not on general philosophical grounds, but on the basis of the way they actually prove in practice their value as guides to research, organizers of findings and mediators of applications.

B. The Problem of Generality

Issues bearing on the suitable level of generality for cognitive theories have been less conspicuous in the literature than those having to do with objectivity, but they have been no less vexing for individual investigators, who must decide how to design and interpret their research. Should the laws and principles of learning and memory apply similarly to all organisms, except perhaps with variations in parameter values, or should different laws apply to different species and different ages within a species? Should cognitive psychology be only the study of a cluster of processes present solely in mature human beings as a consequence of their mastery of language or should it be a study of principles of learning and memory as a manifestation of processes common to both animal and man, young and old, perhaps with elaboration as we go up either scale?

The entire array of conceptual systems—association theory, functionalism, and behavior theory—which dominated research on both human and

animal learning over the first half of the century had in common a view of a hypothetical ageless organism. The hypothetical individual, of course, increased his capabilities for action as a function of experience, but once past an early stage of neural maturation his processes of learning and memory followed the same laws regardless of age (and indeed in some views almost regardless of species). Consequently, we find nearly all of the formal theories of learning and cognition written without reference to the boundary conditions that are to be assumed relative to characteristics of the learner.

The tendency to theorize in terms of an abstract organism may seem unnecessarily sterile, making cognitive psychology both autistic relative to other disciplines and remote from practical affairs. Still, this theoretical strategy should not be too facilely criticized. With a subject matter of such enormous complexity, simplifications are essential if research and theory are even to get a start. Thus, isolation of manageable fragments has had to have first priority. It is good scientific practice to impose simplifications ruthlessly when necessary—but then to remove the simplifying assumptions and discover where and how one can generalize.

In cognitive psychology, the instigation to begin peeking behind the curtain of the abstract organism has begun to arise under the influence of developments in ethology, cross-cultural research, and linguistics. The consequent emergence of at least incipient trends toward theorizing that take account of age and species differences is timely but in a sense by no means new. Rather, we have almost come a full cycle. In the earlier period of the study of learning and animal behavior, it was widely assumed with little justification that a great part of animal behavior was instinctive. Much theoretical and empirical effort was required to clarify the extent to which the original concept of instinct was a tautology and to begin to elucidate the extent to which apparently innate behavior patterns were shaped at least in part through learning processes. But, as usual, the pendulum swung too far, and we saw the development of learning theories that put forth general laws with no attention to the boundary conditions of their applicability with respect either to age or to phylogenetic level of the organism. Then the way back to a more balanced frame of reference had again to be earned by years of arduous research in which findings emerging from the field studies of the ethologists and the laboratory studies of comparative experimental psychologists (for example, Breland & Breland, 1966; Garcia, McGowan, & Green, 1972) brought home the fact that the raw material for the processes described in general laws of learning is not an initial state of subidiocy in which the elements of the nervous system are largely connected at random, but rather an elaborate organization characteristic of a species.

Much the same sequence occurred at the level of human learning, the current corrective influences arising jointly from a vast increase in information concerning linguistic development and the belated appearance of research efforts dealing with human learning in widely different cultural settings (Cole, Gay, Glick, & Sharp, 1971; Cole & Scribner, 1974). Here again, inputs from without were complemented by developments from within the discipline, in this case the formulation of the conception of "control processes" (Atkinson & Shiffrin, 1968) in human memory and the recognition that learned voluntary strategies play a major part in virtually all aspects of human learning.

All of these developments seem salutary for the future of cognitive psychology, the one hazard being that of overcompensation, evidently an almost inevitable characteristic of all intellectual revolutions. It is essential to keep in mind that even the most sophisticated human behavior is more than a collection of voluntary strategies and that understanding behavior requires more than description of strategies and control processes. We need also to understand how strategies and control processes are developed and maintained and how structural constraints impose capacity limitations upon the possible output of control processes. We may find rewards in sacrificing some of the excitement of oscillating between extreme viewpoints in order to take advantage of potentialities of a broader theoretical viewpoint that allows for both general and specific principles of learning and memory and seeks to understand how general principles work their effects within the boundary conditions imposed by particular cases.

With these considerations in mind, we can form some expectations as to how the array of theories of cognition and learning can be ordered on the dimension of generality–specificity of laws or principles. Under the influence of ethology, cross-cultural investigations, and other aspects of the comparative approach (Medin & Cole, Chapter 4 of this volume) we fully appreciate the wealth of differences between species, between age groups within species, between individuals of the same species and age. But there must also be general principles cutting across these categories, for the various types of organisms with their many biological communalities have had to adapt to a common environment, have been subject to common evolutionary pressures, and in the case of human beings have often grown up in common cultural settings. Thus, in analyzing cognitive functions we must expect to find common themes with increasing variation on the themes as we look more closely at particular activities of more narrowly defined classes of organisms.

Further, we can anticipate where in the science of learning general principles will appear: not in the description of observable behavior—there we find endless differentiation—rather in explanatory mechanisms and

principles inferred from behavior, more generality appearing the higher the level of abstraction and the greater the remoteness from direct observation.

We can already see this principle operating even in the most cursory scanning of the various branches of cognitive theory. All organisms, regardless of species or age, must to some extent adapt their behavior to the immediate demands of their environments. They tend to repeat and maintain behaviors that yield satisfaction of needs or that avoid danger. Thus, from the earliest rudimentary stirrings of learning theory we have had principles of hedonism, later the Law of Effect, and still later operant conditioning. These principles, however, are little more than blank checks. The ways in which adjustments are achieved are very different for fish, birds, subhuman mammals, the human infant, the human adult. There are similarities in the forms of predictive laws, but the terms are units of different orders ranging from the simplest reflexes in the case of the lower organisms to behavior chains to highly organized and planned action routines in the case of the human adult.

It seems clear that a full description of the behavior of any organism at any level can, in principle, be given in stimulus–response terms (see, for example, Suppes, 1969). The tendency to perform action X in situation Y is modified in a certain way by outcome Z. If we can specify the environment sufficiently fully, there is little doubt that principles of operant conditioning built on this framework can enable us to predict and control the behavior of any organism to almost any desired degree. Demonstrations of this ability are seen routinely in the cage of the animal trainer, in custodial institutions for the mentally retarded, to some extent even in prisons and reformatories (now the occasion of major public controversy over ethical aspects of "behavior modification").

Why go beyond the progressive refinement of prediction and control in situations where these can be achieved? One obvious answer is the curiosity of some investigators as to the additional principles needed to account for behavior of organisms outside of highly restricted environments. Constantly varying environmental challenges are one of the most conspicuous factors in the evolution of species. We may miss much of importance if we simplify major problems out of existence. Further, some investigators are concerned more with *why* than *how*. Rather than, "How can I rig things to get this organism to do what I wish?," one may ask, "Why do different organisms arrive at such different routes to similar goals?" or, "Why does behavior so often deviate from the optimal which should be so readily obtainable?"

When we persist in asking these questions, we find ourselves forced to begin going beyond the observation of uniformity to the making of inferences, the imagining of uniformities of processes that lie behind variations

in behavior. In a most conspicuous class of instances, we find that to account for differences in behavior in apparently constant situations within or between organisms, we are led to assume that response to the same stimulus varies because of differences in the organism's state of knowledge regarding its environment and differences in expectancies regarding consequences of possible actions. Thus, we are led straight into the psychology of memory—and the heart of cognitive psychology.

The reason why laws and principles of memory are more abstract and more general than those of such forms of learning as operant conditioning is not simply that memory represents an inference from behavior—inferences occur in the interpretations of the simplest conditioning experiments. But there is a difference in the type of inference involved, which goes to the heart of the distinction between learning and memory. In much of the literature of psychology, the terms *learning, retention,* and *memory* are used more or less interchangeably and without attention to niceties of definition. But this practice, though common, does not make for clarity. Looking more closely, we find that there are actually some important uniformities of usage and that recognition of these can help us to understand the different kinds of theories that have arisen.

A distinction of fundamental importance is that *learning* always refers to some systematic change in behavior or behavioral disposition that occurs as a consequence of experience in some specified situation. The learning is a property of a system that comprises an organism together with some limited aspect of the environment within which there prevails some regularities, either deterministic or statistical, which are critical to the organism's adaptation to its immediate circumstances. The appearance of food or the avoidance of some traumatic event is contingent in a specific way upon the organism's actions; or, in the more characteristically human learning situation, a symbolic reward is contingent upon the individual's adjusting his behavior to conform to some criterion of learning imposed by the experimenter or observer. Consequently, the first-order laws of learning are system concepts expressing the way in which the organism modifies its behavior so as to achieve a given end or to approximate a given criterion over some period of time or sequence of trials.

Closely reasoned definitions and applications of system concepts of learning as they arise in the interpretation of classical and operant conditioning, respectively, will be found in the chapters by Kimmel and Burns and by Staddon in Volume 2 of the *Handbook,* and extensions of the same type of analysis to problems of human learning and motivated behavior in chapters by Nelson, Nuttin, and Weiner in Volume 3. To give just one especially clear-cut example of a system concept, we might cite the "matching law," according to which an animal in a choice situation tends to match its rate of responding under any two reward schedules to the relative rates

of reward input yielded by the two schedules (Herrnstein, 1970; Shimp, Volume 2 of this *Handbook,* Chapter 6).

Although it is not in order to go into much more detail here, we should note further that system concepts branch into two major categories according to the time scale over which the principles operate. Instances mentioned in the previous paragraph all have to do with principles describing the way in which an organism adjusts its behavior over some appreciable length of time or series of trials under a given schedule or environmental routine. At the next higher level of abstraction, system concepts refer to the way in which the organism's response tendencies are modified at a particular time or on a particular trial as a consequence of some local combination of events (for example, a stimulus and response followed by a reward). Concepts of the first type may be termed functional descriptions or integrative mechanisms, whereas the second type fit the common conception of causal mechanisms. In the case of classical conditioning, system concepts at the level of causal mechanisms are developed by Rescorla and in the case of operant conditioning by Shimp in Volume 2 of the *Handbook.*

The distinguishing property of concepts of *memory* is that they refer not to any particular organism–environment system but rather to some property or state of the organism which is assumed to have resulted from some experience and which has the consequence of altering the organism's potentialities for a response in any of an infinite number of new test situations to which it might be exposed. This aspect of the concept of memory is generally taken for granted at the level of human behavior, at least outside of the laboratory. We feel, no doubt, that we know what we mean when we speak of an individual's having memory for an event even though we cannot at the time specify or even forsee the future situation in which he may manifest his knowledge. And the same is true for a present-day theorist in the area of memory who speaks of the organization of an individual's memory for a list of words or a set of propositions.

But although memory is unquestionably a higher-order abstraction from behavior, it does not follow that principles of memory are necessarily more general than principles of learning or principles of performance. Rather we should expect to find that principles describing how the memory processes operate should be quite general, but those referring to the content of what is remembered may be quite specific to a particular organism or type of organism with a particular history or type of history. An example of the former class would be the principles of fluctuation of memories between active and passive states over time which appear to operate in a quantitatively simple fashion in situations ranging from classical conditioning to adult verbal memory (Estes, 1955, 1971; Murdock, 1974). Systematic treatments of important concepts in this category with reference to

human memory are developed in chapters by Wickelgren, Bjork & Glenberg, and Craik & Levy in Volume 4 of the *Handbook*.

But although the way in which memories are established and maintained, given that an individual has been subject to a specific sequence of experiences, can be handled by principles of great generality, the same is most emphatically not true concerning the way in which an individual retrieves memories on particular occasions. Whereas the storage of information in memory is an exceedingly abstract concept, the retrieval of information in a test situation is a specific type of performance. Thus memory retrieval must be expected to depend strongly upon particular characteristics of test situations and particular histories of individual organisms. In consequence we find in the extensive treatments of memory in later volumes of the *Handbook* that we can understand a good deal about memory by application of highly abstract, often mathematical, theories of great generality. However, we can do almost nothing toward predicting memory retrieval except by reference to a large body of principles and information, taking account of such factors as species, age, experience, motivation, and task orientations. Although models for memory storage can be, and are, relatively abstract and self-contained, theories and principles of retrieval must be interfaced with theories of learning and motivation and interpreted in terms of a background of knowledge of cognitive development and of the social environment in which the individual is functioning.

It is for these reasons that, even in summary, memory cannot be treated in a single chapter, nor even in a single volume. All I have hoped to achieve here by way of preparation is to make it clear that the enormous variation in treatments of learning and memory that will be encountered results not from whims on the part of various investigators, but from the endlessly complex and multidimensional character of our subject matter.

C. Theory and Application

Like his fellows in all of the other sciences, the investigator of cognition must find a viable position with regard to the question of pure versus applied research. The question is not really whether research should or should not be relevant to real-life problems—no one wants his activities to be basically trivial—but whether relevance must be immediate or can be remote. And when we do feel a need for direct relevance to practical problems, how can the desire be best translated into results: by episodes in which crash programs and research are directed solely to immediately practical problems, or by long-term research programs that include continual attention to more remote objectives?

In order to understand what the present-day investigator of cognitive psychology is doing and why, we need to understand the decisions that

are being made with regard to these general issues and the way in which the results of the decisions are working out in practice.

One of the major classes of problems to which cognitive psychology should be expected to contribute is that of the acquisition and utilization of the basic intellectual skills of everyday life: the elementary uses of language to convey information concerning events, the comprehension of rules and instructions. For the most part these skills are so ubiquitous and acquired so early as to be taken for granted, but serious problems arise when they fail to appear on schedule in mentally retarded children or when they begin to lapse in the senescent period. What can be done to remedy the retarded development or the loss of functions in these cases?

Undoubtedly the causes are often at least partly genetic and perhaps always at least in part biochemical in nature. Thus, there is constant demand for intensified research bearing on the genetic and biochemical causes of elementary intellectual malfunctions. There is every reason to prosecute this kind of research, but little reason to expect that this kind of research in itself can solve problems of mental retardation or disability.

Uncovering the genetic or biochemical differences between normally and abnormally functioning individuals would not in itself solve any practical problems. Physiological and biochemical mechanisms manifest their effects in the organized behavior of the individual, which is in turn a resultant of a long period of interaction between structure and function. In order to be in a position to take advantage of possible advances at the level of genetic or biochemical research, we require a body of theory explaining how intellecutal skills and habits take shape as a joint function of experiential and neurophysiological factors, how the normal course of ontogeny can be modified or reversed by defects at the physiological and biochemical level, how the total intellectual system compensates for defects in one aspect by developing alternative mechanisms to subserve missing or degraded functions.

A second major area of application is that of the elementary educational routines, the teaching of reading, calculating, arithmetical reasoning. Reading is perhaps currently the subject of most attention in view of the perennial wars on illiteracy. Since here our major concern is with the learning of reading and with failures of the learning process, why should we not concentrate our efforts on the early stages of acquisition, preferably focusing on children of the age at which the learning of reading normally begins?

There are many reasons why such a strategy would not necessarily pay off. We could, to be sure, conduct massive studies of the way in which reading is customarily learned. But the nature of the reading process in the child who plods through sentences articulating each syllable and each word as he goes is very different from the performance of the skilled adult reader. Perhaps the performance of the child is a necessary stage, and we should concentrate first on improving our present methods of producing

this performance. But, on the other hand, perhaps the performance actually characteristic of most children while learning to read is not necessary at all. Possibly more efficient instructional methods would take the child through a different succession of stages and one which would not yield such a large harvest of disabilities.

As in other aspects of cognitive psychology having to do with language, we are beginning to appreciate the dangers of devoting too much effort to describing the way in which learning occurs before we know what is learned. Major progress in understanding reading may be unattainable through crash studies of the way reading is customarily learned but may have to wait on the achievement of deeper understanding of the reading process in all its aspects.

Or take the quantitative skills, simple calculation and solving of arithmetical problems. Everyone is familiar with the normal sequence in which these skills are taught. And as the normal sequence is laid out in elementary primers, the order seems so logical that there might seem to be no room for improvement. How could children fail to learn except through laziness or defective nervous systems? The almost ubiquitous failures of our school systems to produce acceptable levels of performance in arithmetical skills on the part of apparently normal children suggests that one might at least have occasion to question whether the customary instructional sequence bears any relation to one that would be optimal from the standpoint of an adequate cognitive theory. Would progress in arithmetic be faster if children started, not with learning of rules for performing addition and subtraction but with studies designed to acquaint them with properties of the number system and relationships between manipulations of numbers and manipulations of sets of objects? To begin to deal with such questions, we need, not comparisons of teaching methods, but a deeper understanding of what has been learned when an individual arrives at various levels of competence in arithmetic and mathematics. In the case of quantitative skills, as in that of reading, major advances in education may well have to wait upon our achievement of deeper understanding of the cognitive processes which the child brings to bear on the tasks, and the way in which these processes come to be organized as a function of different kinds of experience.

IV. THEORETICAL APPROACHES

A. Learning and Behavior Theory

Over the first quarter of this century, the development of theory in psychology was largely confined to the areas of sensory processes and mental abilities. Research on both animal and human learning proceeded within an

eclectic framework and with little obvious acceleration. The first great expansion of research activity was associated with the emergence of the learning and behavior theories of Tolman (1932), Skinner (1938), Lewin (1936), and Hull (1943). This wave of activity channeled a vastly increased proportion of the research effort of psychologists into the study of learning, but the enterprise had little of the flavor that we now associate with cognitive psychology. The theories, and perforce the research, concentrated heavily on the analysis of animal training, and in fact upon only the learning of a few species in a very few experimental paradigms, predominantly the rat in the maze or lever box.

Why this preoccupation with the less intellectual aspects of learning and with rats rather than with man? I think several factors converged to yield this result. Perhaps most important on the theoretical side is the influence of Pavlov and the concept of conditioning. The emergence of the conditioned reflex from the physiological laboratory sparked the hope that basic behavioral units, the building blocks of complex psychological processes, could be abstracted from simple preparations. Further, Pavlov's interpretation of phenomena of conditioning in terms of cortical processes encouraged the idea that the choice of a particular behavior for study is essentially arbitrary—one is as good as another if they all share the same underlying processes in the brain. This orientation resonated perfectly with the increasingly dominant behavioristic philosophy and the powerful influence of operationism on the formation of behavior theories.

But the principal determinants of learning theory were by no means all theoretical, nor even philosophical. The social scene of the 1930s and 1940s saw a vast increase in concern with problems of mental health, the popularization of Freudian doctrines, and the emergence of clinical psychology. For a time, the principal intellectual development in response to expanding needs for psychotherapy was psychoanalysis. Whatever the merits and limitations of psychoanalytic techniques, they were too passive to fit the American lifestyle. Psychologists whose ideas matured in the climate set by Watson and Thorndike could not be satisfied to deal with mental and behavioral disorders by listening to people with problems try to recall childhood experiences. The notion of actively reshaping behavior was far more congenial.

For those eager to get on with the job of reshaping maladaptive behavior, the research on human learning and memory that characterized the period from Ebbinghaus to Thorndike seemed to have little to offer. Studies of human learners dealt always with tiny fragments of activities, as memorization of lists of words, isolated from emotional and motivational processes. In contrast, studies of learning in the chick or the rat might be dealing with relatively simple organisms, but at least they came closer to treating the whole organism and explaining how cognitive and motivational

processes combine to determine behavior. The leading spirits of the age of behavior theory, most especially Skinner aid Hull, clearly articulated the goal of a body of general theory that could be brought to bear on human problems, and they offered routes to this goal by way of specific simplified experimental situations which might be expected to give rise to the needed general principles.

The program of the behavior theorists proved enormously fruitful in some respects but self-limiting in others. Inspired by Hull's vision of a rational quantitative science of behavior, students of learning were for the first time led on a major scale to educate themselves in the philosophy and methodology of formal theory construction. The momentum of this effort is still apparent in the wide ramification of mathematical models for learning and memory.

Even more conspicuously, the ideas and techniques of behavior theory were made to order for investigators seeking either a rapprochement between psychoanalysis and psychology (for example, Dollard, Doob, Miller, Mowrer, & Sears, 1939; Sears, 1943) or a new type of behavioral engineering flowing directly from laboratory to application. The first of these lines of development eventuated in the melding of learning theory and psychotherapy (Dollard & Miller, 1950) and, more recently, the attempts to apply "biofeedback" to the alleviation of physiological disorders (Miller, 1974). The second line led to the body of doctrine and method termed "behavior modification," often controversial but also often fruitful in application to problems of mental retardation and behavior disorders. (Kanfer & Phillips, 1970; Krasner & Ullman, 1965).

Without depreciating its oftentimes impressive accomplishments, we must recognize that behavior theory did not lead in any direct way into the body of theory we find sparking and directing the present broad expansion of research in human cognitive psychology. The reason, in part, seems to be that some of the basic tenets of behavior theory caused it to turn inward to the continual refinement of a particular ensemble of concepts rather than outward to make contact with other strands of theoretical development in psychology.

These remarks are not, I hasten to add, meant to imply that anything went wrong in the fashioning of behavior theory. But rather, it seems a universal characteristic of the scientific enterprise that the mining of any vein of research and theory, however enthusiastically initiated and however diligently prosecuted, tends to go deeper rather than broader and ultimately to become isolated unless diverted into new directions by outside influences. The same was certainly true of research on human learning and memory, largely limited for more than a half century to a few varieties of list memorization initiated by Ebbinghaus (1885) and the simple trial-and-error paradigm of Thorndike (1931). But in the human case, there

was for a long period no major theoretical effort and consequently no substantial acceleration of research activity. The investigation of human memory continued in its relatively placid course, confined by the severely simple conceptual structure of association theory, until the late 1950s, when it was jolted from its narrow path by a succession of intellectual invasions arising from developments in information and communication theory, linguistics, and perhaps most importantly the sciences of computing and information processing.

B. Information Theory and Psycholinguistics

Progress toward general theories of learning and memory had been hampered since the inception of research in these areas by the lack of any adequate means of characterizing what is learned or what is remembered. Efforts to find basic units of experience through introspection proved unrewarding and gave way to the observation and recording of relationships between stimuli and responses. This stimulus–response framework was well suited to the spirit of operationism and seemed in principle to be an adequate basis for the analysis of any one learning situation, provided only that the investigator could decide in advance just what changes in response dispositions would be taken as indices of learning.

But stimulus–response relations, however selected, are necessarily situation specific. Some means were needed to express the consequences of a learning experience in more abstract terms. The need seemed to be met with the appearance of a mathematical theory of communication that offered a means of measuring information in abstract units. The theory was presented in a form accessible to psychologists by Shannon & Weaver (1949), and its implications for the interpretations of cognitive processes, ranging from perception to language, were swiftly exploited in the pacemaking studies of Broadbent (1958), Hovland (1952), and Miller (1956).

As the attention of investigators of human learning began to shift from the modification of response tendencies to the acquisition of information, they began to see ever more clearly that it is essential to distinguish between memory for a string of words and memory for a sentence. An individual who comprehends a printed or spoken sentence may be able to show in a variety of ways that he retains the message it conveyed, whether or not he can recall the particular words. Neither stimulus–response psychology nor any other variety of psychology possessed the conceptual tools to deal constructively with such problems.

But the psychologists of the 1950s discovered that someone else, the linguist, did possess these tools—and the discovery has led to far-reaching changes both in how cognitive psychologists do research and in how they interpret their findings. The direct effects of the influx of linguistic ideas

and methods upon research and theory were substantial (Bower, Chapter 2 of this volume), but an indirect effect may prove to be of even greater moment. I refer to the role of the awakening appreciation of the function of language as an organized system in bridging the gap between the traditionally adult-oriented theories of learning and memory and the burgeoning field of cognitive development. Investigators began to understand the extent to which all cognitive activity is played out against the backdrop of the individual's command of language. Once sensitized to this fact, we find nothing so conspicuous among phenomena of learning and memory as the radical shifts in modes of cognitive functioning that accompany the growing child's increasing competence in the utilization of language (Kendler & Kendler, Chapter 6 of this volume; White, 1965).

Beyond its role in catalyzing the long overdue confluence of learning theory and research on cognitive development, the new discipline of psycholinguistics has provided motivation and techniques for much more rapid and extensive exploration of complex mental processes that would have seemed within reach even a few years ago. It seems certain, however, that the escalation of research on complex linguistic processes would have far outrun our capacity for theoretical assimilation of the research output had it not been for still another major influx of ideas and technology from outside of psychology, in this case originating in the computer revolution.

C. Computer Analogies, State Models, and Simulation Models

Of the extraneous influences that combined to energize the current rise of cognitive psychology, perhaps the most powerful has been the advent of large-scale digital computers, computer science, and computer-based information processing systems. At the most mundane level, the computer simply increased by an order of magnitude our capabilities for handling and analyzing data, thus playing a major, although perhaps not always beneficial, role in the almost exponential increase in volume of research output (the reservation being the hazard that the investigator may find himself increasingly remote from his data and tempted to rely too heavily on packaged statistical routines, thus getting packaged results rather than insights). But in addition to the increase in the volume of output, we have an expansion of the kinds of analyses that are feasible, multidimensional scaling, conjoint measurement, cluster analysis. Of more fundamental significance, we find in our hands a powerful new technique of theory construction—the computer-simulation model.

Behavioral scientists had long chafed under their restriction to the often painfully slow and nearly always arduous task of working out implications of theoretical assumptions regarding complex processes via the classical

methods of mathematics—methods suited to the deep analysis of the behavior of one or a few variables rather than the interactions of many. Presumably, symbolic logic is free of such limitations, but the few serious attempts at application in psychology, for example, Hull, Hovland, Ross, Hall, Perkins, & Fitch (1940), had proved anything but encouraging.

The new art of computer programming placed in the cognitive psychologist's hands a new tool of such unprecedented scope and versatility as to sweep away all technical hindrances to the construction of formal models for complex mental processes, not only those manifest in simplified laboratory tasks but even the acquisition and utilization of natural language (Anderson, 1975; Winograd, 1972). Whether computer simulation models can and will come to be subjected to the same discipline that has made mathematical models such invaluable aids to research in other areas is an open question (Estes, 1975). But perhaps even more important than the body of methods and heuristics contributed by computers and computer programming is the role of the computer-based information-processing system as a new source of analogies and a new theoretical framework for investigations concerned with the structure and function of human memory.

In the early period of research on human learning, theoretical tactics and objectives were strongly influenced by the then new picture of the brain and nervous system as a network of conducting pathways connecting receptors with effectors (see, e.g., Hilgard & Bower, 1966, Chapter 2). It seemed that learning and memory must be based on structural changes in these pathways, hence the conception of stimulus-response connections or associations. The connectionistic model was well suited to the highly analytic orientation of early experimental psychology and structured the pioneering efforts to bring human learning into the laboratory.

However, no theory can thrive indefinitely without competition, and despite the episodes of rebellion on the part of the Gestalt and field theorists, the connectionistic model remained without any serious competitor until the mid-1950s. Then alternatives appeared from two quarters: probability theory and computer science.

First, the flourishing of stochastic models led to a conception of memory in terms of states rather than of connections or associations and such terms as short and long term memory states quickly came into general usage (Atkinson & Crothers, 1964; Bower & Theios, 1964). These models were based on matrix methods, which lent themselves admirably to exploitation by means of digital computers. However, the conception of learning as the passage of the cognitive system through a succession of informational states, like the developmental stages of Piaget, had the one drawback of being in a sense inherently static. The idea that the effects of an individual's past experience relative to a given task could be subsumed in a characterization of his present state is certainly sound. But the conception of a state

contains no explanation of the causal chain of events that takes the learner from one state to the next, no source of insight as to the kinds of observations or experiments that may help unravel the dynamics of the system.

Both the advantages and the limitations of the state models were neatly complemented by the computer-simulation models that appeared on the scene almost simultaneously. The same investigators who first appreciated the power of the digital computer as a research tool in psychology found inspiration also in the rather striking similarities between aspects of a computer-based information-processing system and aspects of the behavior of a human learner or problem solver. Why could one not build a theory of human cognition on the premise that the machinery underlying observed performance resembles, not a network of connections among stimulus and response elements, but rather the internal workings of a digital computer? From this idea it was a short step to flow charts diagramming the memory structure and the succession of information processing operations that might provide accounts of cognitive activities ranging from simple verbal learning (Feigenbaum, 1963) to concept formation (Hunt, 1962) to problem solving (Newell & Simon, 1963).

It is early to evaluate the theoretical advances achieved by the information processing models, but at the least these models have exerted a potent energizing influence on the field. Further, as will be amply documented in this *Handbook,* they have generated an effusion of new research techniques suited for the exploration of cognitive activities more complex and more representative of the uses of language and thought outside the laboratory than might ever have been considered open to investigation within the framework of earlier theories.

D. Interaction of Learning Theory and Cognitive Theory

For those in a position to witness or participate in development of a scientific discipline during a period when it is subject to an infusion of new ideas, the ramifications can be fascinating to observe. The immediate consequence of the impact of the computer revolution was the formulation of a cluster of new models. But it would be quite misleading to imagine that the development of cognitive theory was henceforth confined to information-flow models while association and behavior theories remained frozen in their original posture like petrified trees. The actual course of events has been one of mutual interaction. Association theory has been profoundly altered in orientation and structure under the influence of information-processing concepts. New models emerging from the interaction bear little surface resemblance to those of Ebbinghaus and Thorndike (a notable example being the model of Anderson & Bower, 1973).

Even conditioning and behavior theory have not remained unaffected, as will be seen in Chapter 7 by Bolles in this volume, and those of Shimp in Volume 2 and Medin in Volume 3 of this *Handbook*. It begins to appear that there may be real profit in complementing the almost exclusively behavioral orientation of earlier theories of conditioning and animal learning with one which treats the same phenomena in terms of the processing and transformation of information. The potential advance lies not so much in new explanations of old facts, nor even in the possibility that research on animal learning will be diverted from traditional lines into fruitful new channels, as in the likelihood that informational interpretations may help to maintain and extend lines of communication and influence between the otherwise rapidly diverging fields of human and animal learning.

It is easy for enthusiastic adherents of both disciplines to lose sight of the importance of maintaining intercommunication. Up to a point it is perhaps most efficient to prosecute a given line of research and theory autonomously, without regard for possibilities of reduction to lower or extension to higher levels of analysis. However, the long-term objectives of most investigators of conditioning and animal behavior include possible applications to problems of human behavior. The likelihood of these applications materializing is inversely related to the extent of the intellectual gap that must be crossed.

On the other side, the new wave of research evolving from linguistic and information-processing models may seem far removed from behavior theory. Nonetheless, the individual whose verbal and symbolic activities are under study is an organism as well as an information processor. His behavior, however complex, is a product of the same evolutionary processes that have shaped the behavioral substrate of all living organisms. Although it may be useful to study fragments of his cognitive activity in isolation, the results must be interpreted in relation to other aspects of the functioning of the individual, both biological and social, if they are to be significant for problems arising outside the laboratory.

V. IMPLEMENTING A MULTIDISCIPLINARY STRATEGY

Granted that the investigator of human learning and cognition arrives at an adequate appreciation of the many-faceted character of the subject with which he deals, how should his research activity be affected? At a minimum one might hope for some dampening of the common tendency to succumb to persuasive arguments that only linguistics or only computers or only some other idea or gadget offers the sole answer to the problems of cognitive psychology. The investigator with a broadened background may feel

himself part of a genuinely multidisciplinary enterprise in which acquaintance with a wide variety of concepts and methods is essential if his research is to have impact beyond some very narrow subplot of the field. However, he need by no means become repressed nor inhibited. The dangers of fragmentation of the field into a pandemonium of isolated and competitive subdisciplines arises not from enthusiasm or even overenthusiasm of individuals for their specialties but rather from the tendency to lapse into cultivation of particular research ideas or situations as ends in themselves rather than as avenues to the elucidation of the functioning of the organism in its environment.

The basic reason why we not only should but must be multidisciplinary is that the evolution of the human organism has presented us with a fantastically multilayered system, with old and new portions of the brain, more primitive and more sophisticated learning and behavioral mechanisms all existing and functioning side by side—sometimes in almost total independence, sometimes in intricate interactions. Some of the bodily processes of even the most sophisticated human learner are modified by conditioning in accord with laws that seem indistinguishable from those governing conditioning in lower organisms. In other instances the conditioning processes become overlaid or modulated by verbal and symbolic processes in ways so complex as almost to defy analysis. In some instances the behavioral subsystem that is undergoing conditioning may be sufficiently isolated as to be conveniently ignored by the investigator of verbal or symbolic activities. In other instances, conditioning processes are inextricably involved in the development of emotions, motivations, or attitudes that are most relevant to the way in which the individual's cognitive state influences his behavior.

At the same time, one should not view mechanisms of conditioning and reinforcement solely in terms of their bearing on vegetative and motivational functions, which are largely obscured in the more cognitive activities of the adult human being. Of necessity, research on conditioning in lower organisms has sometimes led into periods of refined and exacting analysis of the manner in which particular response systems are modified, yielding results that can only be interpreted by a few specialists. Nonetheless, research on conditioning in the same organisms, and in fact sometimes in the same experiments, bears on the processes and mechanisms by which animals gain information about their environments. On the surface, these processes are quite remote from those we observe in the adult human being, but for understanding both lower and higher organisms the similarities may be more significant than the differences. Principles of operant conditioning and reinforcement derived from the animal laboratory appear to apply almost without modification to some behavioral changes in the very young

child or in the mentally retarded. Evidence that the same mechanisms are at work can be detected in the adult by means of delicately contrived experiments, but in normal behavior these mechanisms operate only against the background of complex symbolic processes that must be understood in the context of broader theories of memory and cognition (Atkinson & Wickens, 1971; Estes, 1969, 1972).

In many intellectual tasks, and especially in situations which tax his capabilities, the adult human learner functions in a manner that makes him appear less like a sophisticated animal than like a warm blooded computer. Nonetheless, with continuing observation, he also will be found to manifest limitations of capacity and to lapse into disabilities and malfunctions that would not be tolerated in a well engineered computer. To understand why he should be subject to these limitations and these malfunctions one needs to understand him as a product of the evolution of the species and the development of the individual.

Research focused on the description and analysis of the human learner's performance on specific tasks requires increasingly specialized techniques and it often must proceed on a day-to-day basis with attention confined to the immediate demands of technical methodology and immediately relevant theories. But if the results are to contribute to an understanding of the functioning of the individual outside of research situations, we need periodically to consider also the problem of placing the cumulating findings in the broader perspective of the array of theories treating human behavior in relation to both biological and cultural determinants.

The emerging picture of what an investigator of learning and cognition needs to know and needs to do in order to pursue his specialty effectively is a bit staggering. Here perhaps lies a critical but almost untouched area of application for our findings. To borrow once again from computer terminology, we evidently require an Operation Bootstrap—that is, intensive application of what we learn about learning to the problem of the scientist who hopes to keep up with the increasingly stringent demands of this challenging enterprise.

REFERENCES

Anderson, J. R. Computer simulation of a language acquisition system: First report. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1975.

Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D.C.: Winston, 1973.

Atkinson, R. C., & Crothers, E. J. A comparison of paired-associate learning models having different acquisition and retention axioms. *Journal of Mathematical Psychology*, 1964, **2**, 285–315.

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. Pp. 89–195.

Atkinson, R. C., & Wickens, T. D. Human memory and the concept of reinforcement. In R. Glaser (Ed.), *The nature of reinforcement.* New York: Academic Press, 1971. Pp. 66–120.

Bower, G. H., & Theios, J. A learning model for discrete performance levels. In R. C. Atkinson (Ed.), *Studies in mathematical psychology.* Stanford, California: Stanford University Press, 1964. Pp. 1–31.

Breland, K., & Breland, M. *Animal behavior.* New York: Macmillan, 1966.

Broadbent, D. E. *Perception and communication.* New York: Pergamon Press, 1958.

Cole, M., Gay, J., Glick, J., & Sharp, D. W. *The cultural context of learning and thinking.* New York: Basic Books, 1971.

Cole, M., & Scribner, S. *Culture and thought. A psychological introduction.* New York: Wiley, 1974.

Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. *Frustration and aggression.* New Haven, Connecticut: Yale University Press, 1939.

Dollard, J., & Miller, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.

Ebbinghaus, H. *Memory.* (Translated from the 1885 original by H. A. Ruger and C. E. Bussenius.) New York: Teachers College, 1913.

Estes, W. K. Statistical theory of spontaneous recovery and regression. *Psychological Review,* 1955, **62,** 145–154.

Estes, W. K. Reinforcement in human learning. In J. Tapp (Ed.), *Reinforcement and behavior.* New York: Academic Press, 1969. Pp. 63–94.

Estes, W. K. Learning and memory. In E. F. Beckenbach (Ed.), *Concepts of communication.* New York: Wiley, 1971. Pp. 282–300.

Estes, W. K. Reinforcement in human behavior. *American Scientist,* 1972, **60,** 723–729.

Estes, W. K. Some targets for mathematical psychology. *Journal of Mathematical Psychology,* 1975, **12,** in press.

Feigenbaum, E. A. The simulation of verbal learning behavior. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought.* New York: McGraw-Hill, 1963. Pp. 297–309.

Garcia, J., McGowan, B. K., & Green, K. F. Biological constraints on conditioning. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory.* New York: Appleton-Century-Crofts, 1972. Pp. 3–27.

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

Hilgard, E. R., & Bower, G. H. *Theories of learning.* New York: Appleton-Century-Crofts, 1966.

Hovland, C. I. A "communication analysis" of concept learning. *Psychological Review,* 1952, **59,** 461–472.

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

Hull, C. L., Hovland, C. I., Ross, R. T., Hall, M., Perkins, D. T., & Fitch, F. B. *Mathematico-deductive theory of rote learning.* New Haven, Connecticut: Yale University Press, 1940.

Hunt, E. B. *Concept learning: An information processing problem.* New York: Wiley, 1962.

Kanfer, F. H., & Phillips, S. *Learning foundations of behavior therapy.* New York: Wiley, 1970.

Krasner, L., & Ullman, L. P. *Research in behavior modification.* New York: Holt, Rhinehart & Winston, 1965.

Lewin, K. *Principles of topological psychology.* (Translated by F. Heider & G. M. Heider). New York: McGraw-Hill, 1936.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81–97.

Miller, N. E. Applications of learning and biofeedback to psychiatry and medicine. In A. M. Freedman, H. I. Kaplan, & B. J. Sadock (Eds.), *Comprehensive textbook of psychiatry*. (2nd ed.) Baltimore, Maryland: Williams & Wilkins, 1974.

Murdock, B. B., Jr. *Human memory: Theory and data.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1974.

Newell, A., & Simon, H. A. GPS, a program that simulates human thought. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought*. New York: McGraw-Hill, 1963.

Pavlov, I. P. *Conditional reflexes.* (Translated by G. V. Anrep) London: Oxford University Press, 1927. (Reprinted by Dover, New York, 1960.)

Sears, R. R. Survey of objective studies of psychoanalytic concepts. A report prepared for the Committee on Social Adjustment. *Social Science Research Council Bulletin* **51**, 1943.

Shannon, C. E., & Weaver, W. *The mathematical theory of communication.* Urbana, Illinois: University of Illinois Press, 1949.

Skinner, B. F. *The behavior of organisms.* Appleton-Century-Crofts, 1938.

Suppes, P. Stimulus–response theory of automata and TOTE hierarchies: A reply to Arbib. *Psychological Review*, 1969, **76**, 511–514.

Thorndike, E. L. *Human learning.* New York: Century, 1931.

Titchener, E. B. *Experimental psychology. Qualitative.* New York: Macmillan, 1909.

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

Watson, J. B. *Behaviorism.* New York: W. W. Norton, 1924.

White, S. H. Evidence for a hierarchical arrangement of learning processes. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in child development and behavior.* New York: Academic Press, 1965. Pp. 187–220.

Winograd, T. Understanding natural language. *Cognitive Psychology*, 1972, **3**, 1–72.

2

Cognitive Psychology: An Introduction

Gordon H. Bower

Stanford University

I. INTRODUCTION

Cognitive psychology is concerned with how organisms cognize or gain knowledge about their world, and how they use that knowledge to guide decisions and perform effective actions. Knowledge enables us to survive in a hostile environment, to satisfy our social and biological needs, to plan for our own and our children's future. Organisms exhibit several broad classes of knowledge: knowledge that certain propositions are true (so-called factual knowledge), and knowledge of how to do a variety of things (skills, procedures). Historically, stimulus–response psychology emphasized motor skills, whereas cognitive psychology emphasized factual knowledge. But obviously an organism needs both types of knowledge to get along in this world. Because the cognitive system has developed as an instrument of adjustment, as an aid to satisfying needs and motives, it is appropriate that we begin with discussion of biological and social motives, and their directive function.

A. Motives Driving the Cognitive System

1. Biological Motives

We have evolved as biological machines under the selective pressures of the "contingencies of survival," as Skinner (1974) calls them. Evolution eliminated those species that did not have adaptive exchanges with their environments. Body cells require periodic replenishing of their biochemical

nutrients; organisms that interact effectively with their environment to get those nutrients are the ones that survive.

A set of interconnected homeostatic mechanisms has evolved to handle the biological emergencies arising when the equilibrium of the body's internal milieu is disturbed. For example, when a person becomes hot, this condition is detected by heat-sensitive cells in the vascular bed of the hypothalamus. The input to those sensors passes some threshold which switches on several vasomotor reflexes that will produce bodily cooling: the capillaries dilate, carrying more red blood to the skin surface for cooling; the skin sweats, so that evaporation of the water produces a cooling effect; and the organism is impelled (motivated) to carry out actions whose ultimate goal is cooling—he sheds excess clothing, he stands in a breeze or in the shade, he seeks an air-conditioned room, or he may say, "Would you please open the window?" etc. All such activities are instrumental means to the end of reestablishing homeostasis in the internal milieu.

Knowledge and motor skills play major roles in this adaptive process. First, as a precondition, the adaptive organism must be able to identify internal states (for example, cooling of the body core) that comprise favorable changes in the internal milieu. Second, through associative learning similar to Pavlovian conditioning, he must learn to identify external situations or happenings in the world which will bring about (or are at least correlated with) favorable internal changes. These situations (for example, a cool room) are subgoals, and they come to be valued as a means to a desired end. Third, the organism must be able to acquire and then apply coordinated action patterns that bring about these valued goals. This is the "situation–action–outcome" knowledge acquired through the process we call operant conditioning. For example, a person must learn how to turn on an air conditioner, or how to get to swimming pools and air conditioned rooms that are located nearly. It is particularly in regard to these latter learnings—in effective interactions with the environment—that the cognitive system seems to have developed to the highest degree.

2. Social Motives

We have illustrated the motives underlying behavior in terms of a biological need (temperature regulation), although other biological needs like hunger and thirst would have served as equally cogent illustrations. However, it is clear that organisms pursue many autonomous goals that are not of this impelling metabolic kind. We engage in many activities for their own sake—hiking, painting, building toys, playing music, playing games of all sorts, and so forth. We probably also have "cognitive motives," such

as the desire for conceptual clarity in puzzling situations, the desire to reduce uncertainty about future events, and the desire to know how or why things work. Maslow (1967) proposed that people have an ordered scale of motives, ranging from biological needs at the lowest end of the scale through social needs (such as affiliation, love) to more transcendent personal needs such as self-awareness, actualization, and personal growth at the highest end. The customs of modern societies are partly organized around satisfying automatically the lower motives. Maslow hypothesized that people would pursue motives at a given level only if those at lower levels had been satisfied. Judging from the evidence of social and personal disorganization in concentration camps, Maslow's conjecture is at least plausible.

3. *Dealing with Multiple Motives*

People are obviously guided by multiple motives. Several methods have envolved to deal with this multiplicity (see Simon, 1967). The first method is by "priority queueing" of goals in terms of their urgency: if, while I am pursuing satisfaction of Motive 1 (say, I am a salesman selling a car), conditions arise that give prominence to Motive 2 (say, going to the toilet), then depending on the relative urgencies the cognitive system can either (1) queue up the newly arrived goal in a waiting line to be serviced, or (2) interrupt activity towards Goal 1 in order to turn momentarily to Goal 2, and then return to Goal 1.

A final method available for dealing with multiple goals is to seek out activities which will satisfy several goals simultaneously. An example of a multigoal activity is the "business lunch": it serves the goal of doing business as well as reducing hunger. Furthermore, the exact form, surroundings, and contents of the lunch may satisfy other goals of aesthetics, quiet relaxation, and perhaps enhancement of self-esteem. Sometimes the main factor determining selection among two actions is the greater number of motives that can be satisfied by one rather than the other action.

B. Intentions, Purposes, and Actions

These matters of motives are mentioned at the outset to reveal a fundamental assumption of cognitive psychology, namely, that the unit of behavior is the "purposive action" rather than the "colorless movement" or "glandular squirt" of yesterday's behaviorists. In this regard, cognitive psychologists side with the layman who uses an action terminology committed to the idea that goals explain or account for the behavior. Thus, when asked "*What* is John doing?," we can describe the action by some

remark like, "He's trying to build a fire." But the very same answer also serves as the layman's *explanation* when we ask, "*Why* is John doing that?" In characterizing the action, we are also characterizing the intention or goal of the action, and the layman's theory explains behavior as an expression of intentions, desires, and the like.

Tolman (1932) and other cognitive psychologists thought such explanations were proper though incomplete; he recognized that one had yet to explain why exactly *this* intention had arisen and *this* class of actions had been chosen to achieve the goal—but Tolman supplied ready answers in terms of demands and past learning experiences. The stimulus–response behaviorists such as Watson, Guthrie, and Hull objected because they thought psychology should deal with objective movements—that is, muscle twitches and glandular secretions—and only explain purposive actions in terms of these elements. They thought explanations couched in terms of purposes and goals were teleological and necessarily unscientific.

Extensive philosophical analysis has been devoted to the concepts of actions, movements, intentions, and teleological explanations, (see particularly Taylor, 1964). The upshot, I think, is that the behaviorist's insistance upon colorless, purposeless movements as the "units" of behavior is best viewed as a mistake. Actions are simply not reducible to, or equivalent to, a set of movements. It is practically impossible to describe a response without reference to its achievements, or to the stimuli with which the effectors (muscle groups) are in contact, and to the feedback provided by these effectors. Moreover, a given movement sequence can be two different "actions" depending upon the inferred intention or "meaning" of the action to the actor. This difference is captured in our everyday distinction between "accidental" and "intentional" responses and consequences. If a man hits me in the face as he is stretching, I infer different intentions if I know he is asleep than if I know he can see exactly what he is doing. The distinction also appears throughout the law: for example, a defendant is charged with first-degree murder if the killing was intentional, is charged with manslaughter if the killing was unintentional but related to his "negligent" actions, but is not charged at all if the killing was clearly "accidental."

Finally, action units simplify descriptions of behavior since we aggregate behaviors according to their immanent effect or goal, and this suppresses large amounts of irrelevant variability at the molecular level of movements. Thus, for most theoretical purposes, it is adequate to say that a man is "building a fire," even though it is understood that there are many different ways to build fires depending upon circumstances, each with ideosyncratic variations. Action terms serve as classificatory concepts; because of this, it would not be inappropriate to say that the man is behaving in such manner as to produce or realize an instance of the action concept we call "fire

building." Our procedures for assigning a stretch of behavior to an action class are similar to those by which we assign physical objects and events to classes. A significant difference is that in order to qualify as an instance of action class X, the behavior must vary according to the circumstances in such manner as to achieve X in those circumstances. Another illustration of the open-ended character of action concepts arises with locomotion in space, as when we say, "I'm going to the grocery store." In getting from point A to point B in town, we know and could use multiple routes and means of conveyance as circumstances require: we could go in a wheelchair or crawl or hobble on stilts if required to, even though we have never gone from A to B in any of these ways. This is not to say that once I learn one way to go from A to B, I automatically know all possible ways to get from A to B (which is absurd). Some earlier sensorimotor learning is clearly needed in order to apply or transfer particular skills to the present problem.

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

C. Plans and Hierarchies

1. Plans

This concept of a plan was introduced by Miller, Galanter, and Pribram (1960) and it appears throughout the writings of later psychologists. A plan may be thought of as a procedure or recipe for achieving some goal. Most plans have a hierarchical structure, which proves useful for analyzing large segments of goal-directed behavior. A given goal will generate several subordinate goals (called subgoals) possibly with their own subsubgoals. The logical arrangement of subgoals within goals can be diagramed as in Figure 1, which illustrates a plan that goes two levels deep. Suppose your top goal is to go to a theater play tonight. To do this, you must first reserve tickets (goal A). To reserve tickets, you must first look up the telephone number of the theater (do C), then telephone and reserve tickets (do D). If that succeeds, then this evening you must get to your seat at the theater on time (goal B). This requires leaving on time, driving your car to the theater, and parking (do E), then paying for your tickets and walking to your seat for the start of the play (do F). By satisfying subsubgoal F, you also satisfy subgoal B, which is a precondition for your satisfying

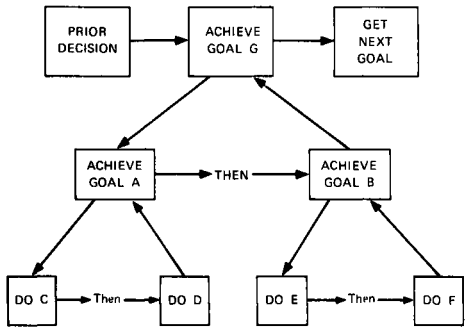

FIGURE 1 Diagram of a hierarchical arrangement of subgoals embedded within a top-level goal.

the top goal *G*. After that, you might begin operating on a new goal, such as finding a coffeehouse for refreshments after the concert.

These kinds of hierarchical analyses of action plans have considerable appeal; they are a neat way of parsing or chunking the significant aspects of behavior. One can refine and expand the behavioral description to as much detail as is of interest to the current investigation. Thus, the subgoal "look up telephone number" is a primitive in the description of Figure 1, but is differentiable in terms of other subgoals and elementary situation–action routines, for example, "find the phonebook," "open to desired page," "scan list for name." Each of these in turn is further differentiable, for example, "scan list" means to look at each entry successively until you find one that matches the name desired.

An attractive feature of this approach is that it solves the ancient "molar versus molecular" dispute concerning the proper level at which behavior should be characterized by the observing psychologist. As is well known, a given bit of behavior can be described at various levels; we could say that our subject is making particular subvocal movements, that he is comparing the name "Bijou" to his target, that he is moving his finger down the page, that he is looking up a phone number, that he is making ticket reservations, that he is planning to go to the theatre, that he is promoting his personal growth, and so on. It is senseless to ask which of these things is he "really" doing; the answer is that he is doing all of them, more or less at the same time. They are simply descriptions that focus on subgoals at different levels of the hierarchical plan he is carrying out.

These hierarchical descriptions of plans are quite congenial to the information-processing approach since plans resemble flow charts showing the passage of control among different components (the "subroutines") of a hierarchical program that is solving a problem. The behavioral plan may thus be represented by us as a computer program, a sequence of instructions and branching decisions for carrying out some actions that are likely to achieve some goal. When the person "executes" a plan, he proceeds through it step by step, completing one part and then moving on to the

next. A very elementary plan for "walking the length of a block" (Simon, 1967, p. 31) would be:

Walk the length of a block:

1. Step with left foot, then 2.
2. Step with right foot, then 3.
3. If end of block, do 4; if not, do 1.
4. Terminate.

This is a list of elementary instructions to be performed in order, with a conditional test inserted at the third line. A similar small program might be written for "how to cross an intersection." By combining these two programs, we can synthesize somewhat larger or more interesting segments of behavior, such as "Walk seven blocks north," thus:

Walk 7 blocks north:

1. Face north, then 2.
2. Set count = 0, then 3.
3. Walk the length of a block, then 4.
4. Add 1 to the count, then 5.
5. If count = 7, do 7; else, do 6.
6. Cross intersection, then 3.
7. Terminate.

The important point is that "walk a block" and "cross intersection," although themselves programs, can be used as units within a larger program. These smaller segments are called subroutines and may be characterized or named in terms of what functions they perform. In today's programs to simulate learning or problem solving, there would be a hierarchy of subroutines that are organized in some way designed to achieve particular goals.

You might ask: *What* reads and executes these alledged instructions in your alleged plans? In the computer, of course, it is the *interpreter* which "reads" instructions and translates them into physical happenings. But when one proposes a program (or mechanical algorithm) to simulate behavior, he need not be proposing that there is some inner interpreter (homunculus) that "reads" the physiological counterpart of the instructions and fires off the effector movements. Rather, the program is just *our description* of the system; it is external to the subject himself. The program need not correspond to anything inside the person; rather, the program may only behave the same way as the mechanisms that do control the behavior. For instance, one could write a program to simulate a thermostat, but he would not thereby be committed to the assertion that the expansion wire in the thermostat is "following" that program. In contrast, there may

be in a real system something that does correspond to a program that causally controls its behavior. An example would be the pattern of holes in the paper player roll that controls the movement of keys in an old-fashioned player piano (the examples come from Newell & Simon, 1972). In this case, the "interpreter" is the set of metal fingers pressing against the player roll that sense the holes and activate, by air pressure, their corresponding keys. Clearly, the interpreter can be quite "mechanical" and is not an occult entity. Whether or not one imputes physiological "reality" to his program, our external knowledge of the program gives us predictive power regarding the behavior of the device it simulates or controls.

2. *The Executive Monitor*

In a hierarchical program, the top-level routine is called the *executive,* and this is a useful concept in models of the cognitive system. The executive calls routines at the next lower level and keeps track of where these subroutines are to return their results. The executive monitors the number of subgoals being generated using a particular method (say, in a problem-solving situation). It also evaluates (from feedback) how the current plan is progressing. The executive may interrupt and switch to another subgoal either if that other one suddenly becomes more important or if the current method of attack on a subgoal seems not to be progressing satisfactory, for example, if it exceeds a work limit. The executive also notices when a subgoal has been completed so that its results may be used in selecting the next step or next goal to be worked on. The executive is itself just another bit of program and is quite deterministic in its actions.

3. *Plans for Learning*

The primary concern of this chapter is with learning and remembering, so we shall usually be discussing the operations of the cognitive system when the current plan controlling activity is either a plan for learning (acquisition of specified information) or a plan for remembering (retrieving previously stored information). In characterizing learning, one is interested in analyzing the means by which the perceptual process enters information into short-term memory, and in how that information comes to be represented and related to other things as it is stored in long-term memory. For example, we may regard memorizing as a problem-solving task for which the person selects and applies particular "mnemonic strategies" or plans which have proven useful to him for solving similar memorization problems. These plans for deliberate memorization (for example, concentration and verbal rehearsal) are clearly learned and develop along side other intellectual abilities of the child (see Meacham, 1972). It is clear

too that the degree of learning we observe depends heavily upon which mnemonic strategy the person employs when he is exposed to the material.

With respect to plans for remembering, the significant issues to be addressed are (1) how the current "test question" or problematic situation can retrieve from memory information relevant to it, (2) what is the organization or "filing index" of the storage system, (3) what search procedures are carried out to locate relevant memories, and (4) how do we construct answers based on the fragmentary information we retrieve.

II. INFORMATION-PROCESSING ANALYSIS

A. Stages and Components

The information-processing approach assumes that perception and learning can be analyzed conceptually into a series of stages during which particular components ("mechanisms") perform certain transformations or recodings of the information coming into them. The subject's eventual response (for example, the perceptual judgment "I see a giraffe") is considered to be the outcome of this lengthy series of operations. Each stage in the system receives as input the information as coded in its predecessor stage, operates upon it so as to condense, abstract, recode, or elaborate it, and then passes this product on to the next stage in the analysis. Since external stimuli can not get inside an organism, the representation of them ("internal symbols") and their interrelations ("symbol structures") is what we call "information," and this is the content we describe in our theories. As experimenters, we try to devise techniques for studying the representation of information at each stage, the nature of the recoding done at that stage, the experimental variables that influence the duration of the stage, and so on. We try to represent the supposed causal order of these processes in terms of a flow diagram in which blocks represent component processes occurring at successive intervals. Each block is labeled according to its typical function. In terms of the computer analogy, the operation of each processing stage would be represented by a subroutine. The identification of these processes and their causal operation remains the major, unfinished task for information-processing theories.

B. Some Concepts of Information Processing

However, before launching into a description of the "human system," it is prudent to mention briefly some of the terminology to be used as we proceed. According to Newell and Simon (1972, p. 20 ff.), an information processing system (IPS) consists of a sensory system, a response generator,

a memory, and a central processor. They then provide the following set of interrelated definitions:

1. There is a set of elements, called *symbols.*
2. A symbol structure consists of a set of *tokens* (equivalently, *instances* or *occurrences*) of symbols connected by a set of *relations.*
3. A *memory* which is a component of an IPS capable of storing and retaining symbol structures.
4. An *information process* is a process that has symbol structures for (some of) its inputs or outputs.
5. A *processor* is a component of an IPS consisting of:
   (a) a (fixed) set of *elementary information processes* (eip's);
   (b) a short-term memory (STM) that holds the input and output symbol structures of the eip's;
   (c) an *interpreter* that determines the sequence of eip's to be executed by the IPS as a function of the symbol structures in STM.
6. A symbol structure *designates* (equivalently, *references* or *points to*) an object if there exist information processes that admit the symbol structure as input and either:
   (a) affect the object, or
   (b) produce, as output, symbol structures that depend on the object.
7. A symbol structure is a *program* if (a) the object it designates is an information process, and (b) the interpreter, if given the program, can execute the designated process. (Literally this should read, "if given an input that designates the program.")
8. A symbol is *primitive* if its designation (or its creation) is fixed by the elementary information processes or by the external environment of the IPS [Newell & Simon, 1972, p. 20].

Also,

> *Reading* or encoding consists in creating in memory internal symbol structures that designate external stimuli; *writing* is the inverse operation of creating responses in the external environment that are designated by internal symbol structures [Newell & Simon, 1972, p. 21].

An illustrative example from Newell and Simon (1972, pp. 21–22) is a program for translating Morse code. Thus, an external stimulus such as (—, —, —) would be read (recognized) in short-term memory as a symbol structure, namely, three tokens of the dash symbol connected by the "next" relation. (The "next" relation is used in building up lists of symbols.) Next, the designation of this symbol structure would be "looked up" in long-term memory. For this example, the corresponding symbol structure in long-term memory is attached to the letter S. Hence, the (—, —, —) structure in short-term memory would be replaced by its designation code, S. As the receiver listens, he builds up an *encoded* symbol structure in short-term memory such as (S,I,T), which itself in a later pass designates a word. This discussion describes stimulus recognition at intake.

The process of sending or writing (generation) would just use the machinery in the inverse order.

C. Elementary Information Processes

There would be a set of elementary information processes that can be executed by the system. Among them would be one to "find the member on a list that is next to a specified one"; another one would "check whether two symbol tokens, x and y, are identical." By combining these two elementary processes, one can synthesize a larger process, such as one which compares two lists of symbols to see whether they are identical.

A "symbol" in such a system is a unit which permits access to any information associated with it: this information might be its "referent" or another symbol structure. For example, the internal symbol "John Smith" might refer to—be associated with—a symbol structure describing his appearance and other facts about him.

A major claim of computer science is that there is a relatively small set of elementary information processes which will suffice to produce the full generality of any symbol-manipulating and problem-solving activity we could want. Newell and Simon suggest that the following ones comprise a sufficient set:

1. *Discrimination.* It must be possible for the system to behave in alternative ways depending on what symbol structures are in its STM. Furthermore, the behavior needs to be *arbitrarily* alterable; i.e., transfer of control to an independent program must be possible.
2. *Tests and comparisons.* It must be possible to determine that two symbol tokens do or do not belong to the same symbol type. Often comparisons are directly coupled with conditional behavior, but they may equally well lead to the production of a conventional symbol (e.g., *true* or *false*) that can later be discriminated.
3. *Symbol creation.* It must be possible to create new symbols and set them to designate specified symbol structures. Again, this process must be performable arbitrarily; i.e., whenever a new symbol is desired it can be created, and it carries no meaning other than that it designates the desired symbol structure. Whether the system must also be able to destroy symbols depends primarily on whether memory capacity is limited.
4. *Writing symbol structures.* It must be possible to create a new symbol structure, copy an existing symbol structure, and modify an existing symbol structure, either by changing or deleting symbol tokens belonging to the structure or by appending new tokens with specified relations to the structure. Many variations are possible, as long as they permit building up arbitrary structures.
5. *Reading and writing externally.* It must be possible to designate stimuli received from the external environment by means of internal symbols or symbol structures, and to produce external responses as a function of internal symbol structures that designate these responses.
6. *Designating symbol structures.* It must be possible to designate various parts of any given symbol structure, and to obtain designations of other

parts, as a function of given parts and relations. Again, this can be achieved in many ways but it must be always possible; i.e., there must not be any parts of symbol structures that are in principle inaccessible.

7. *Storing symbol structures.* It must be possible to remember a symbol structure for later use, by storing it in the memory and retrieving it at any arbitrary time via a symbol structure that designates it. How much memory is available, of course, conditions strongly how complex the totality of stored structures may be. The memory must be highly reliable over time [Newell & Simon, 1972, pp. 29–30].

It is supposed that such components as these will be needed in building a model to simulate some interesting behavior. Assuming the existence of such information processes does not commit one to any particular level of analyzing a given psychological phenomenon. What may be primitive symbols at one level of analysis (for example, words for language understanding programs) are large symbol structures in models aimed at a more molecular level of analysis (for example, words as ordered sets of graphemic features in models of word perception). The content of the above processes will also vary with particular applications—for example, the "external responses" that can be produced (under 5 above) obviously differ depending on whether one is modeling a crayfish, an infant, or a human adult.

Whatever the phenomenon being modeled, it is assumed that most of the above processes will come into play at some level of analysis. To illustrate, in modeling the memory-scanning task of Sternberg (1969), one might assume that presentation of the stimuli "7, 9, 3, 6" as the memory set causes the subject to read these in (Process 5 above), to match them to symbol types in long-term memory (Processes 1 and 2), and set up copies (Process 3) of these symbol types in a symbol structure in short-term memory (Process 7). Thus, a element called "Most recent list" is made to designate the symbol structure "7–next–9–next 3–next–6" (by Processes 3 and 4). The test probe, "Is 3 in the most recent list?," is read in (Process 5), compared symbol by symbol to successive elements on the most recent list (using Processes 1, 2, 6), with a "yes" response being generated (Process 5) because the test probe matches one of the items in the set. The operation of "get the next symbol on the list" is an elementary process (6 above) as is the process of "compare list symbol to probe symbol for a match" (2 above). We thus see how a given model analysis uses such information components as these.

III. OVERVIEW OF THE COGNITIVE SYSTEM

With this introduction to basic vocabulary and concepts of the information processing approach, let us turn to sketching the viewpoint as it applies to the behaving organism. As noted above, the approach tries to represent

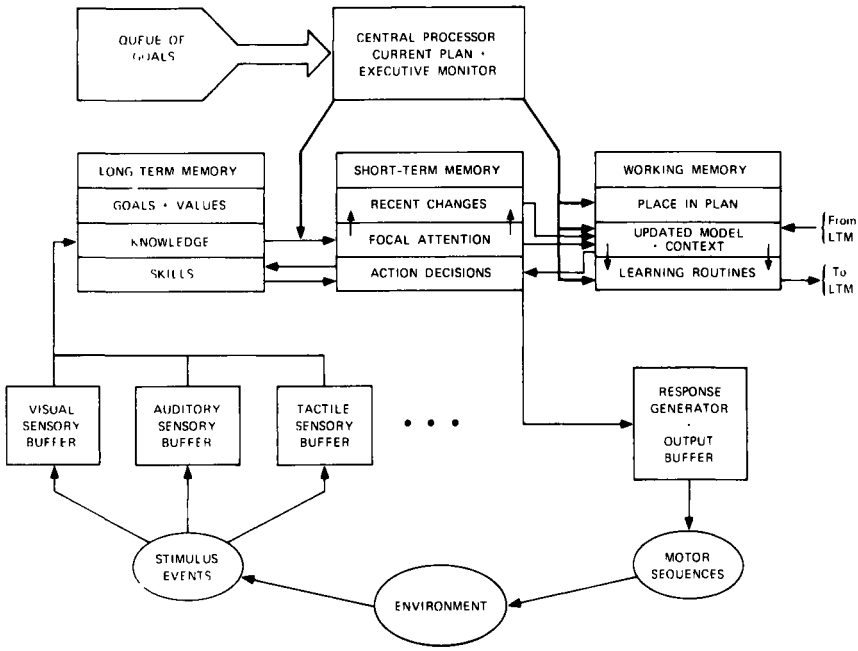

FIGURE 2 Diagram of the principal components of the perception and memory system. Arrows indicate the direction of information flow and control among the components.

the cognitive system in terms of a flow diagram in which blocks represent components and successive stages of processing. Although many such flow charts have been proposed, a sort of "modal" summary chart might look like that in Figure 2. The interconnections between components are specified by the arrows that suggest the direction of action and control. The components may be divided roughly into the sensory system, the response system, the long-term memory system, and the central processor wherein occurs the active processes coordinate with memorizing, thinking, evaluating, and decision making. We now describe these components in Figure 2, starting at the sensory end.

A. Perception

1. Sensory Buffers

The environment provides an array of patterned stimulus energies, reflecting both static settings as well as dynamic changes ("events"). Our sensory receptors have transducers which convert these analog energy patterns into a digital code (that is, nerve firing patterns). Apparently at a very low level in the sensory system (for vision, the retinal ganglion cells), certain elementary "preattentive" processes come into play: in vision

(which will always serve as our illustration), these processes perform contour enhancement, centering of the stimulus, segregation of the figure from the background, and so on. Quite a bit is known about the neural mechanisms involved at this stage (see Ratliff, 1965; von Békésy, 1967).

As the stimulus pattern is registered on the receptor surface (say, the letter H projected on the retina), a set of feature detectors located along the visual pathways begin extracting significant features from the stimulus. Features of H would be that it is black, has two vertical lines, one horizontal line, is open at the top and bottom, etc. These features are apparently extracted in parallel for all patterns falling near the fovea. The feature coding at this level is quite automatic and seems little affected by momentary expectations. The physiology underlying and corresponding to feature detection has been intensively investigated, steming from the classic work of Hubel and Wiesel (1962, 1968) and many others. We now know how to construct plausible neuronal-network models that will extract a large and relatively sophisticated set of features such as the size, orientation, direction of motion, and velocity of movement of lines, edges, arcs, and so on (see, e.g., Lindsay & Norman, 1972, Chapter 2). It appears that the connecting up of the neural elements which extract a given feature (for example, horizontal lines) requires that the developing infant receive normal visual stimulation, since restricting vision during rearing (say, to only vertical lines) causes selective deficits in appropriate cortical detecters in kittens (see review by Blakemore, 1974). Restrictive environments aside, Eleanor Gibson (1969) has argued that perceptual development should be viewed in terms of the organisms's growing sensitivity to more, and more subtle, distinctive features of his perceptual environment.

If the stimulus is a brief change (say, a letter flashed in a tachistoscope), the stimulus is held briefly in a sensory memory as an image or icon while its features are being extracted. We represent this by saying that the image is held in a "visual buffer" as a sort of fleeting memory of the physical stimulus. The duration of the icon can be varied by the field-to-flash contrast before and after the flash. The icon can be "washed out" by a bright masking flash just after the shape is flashed, or it can be replaced by a second form being flashed very soon afterwards to the same retinal location (or to the homologous location on the other retina, see Turvey, 1973). The brief duration of the icon severely limits the number of features which can be extracted from a brief glimpse. In typical conditions the icon lasts about a quarter of a second which, interestingly, is also about the time for a saccadic eye movement. This suggests that as the eye sweeps over a scene, the icon from view $n + 1$ replaces (rather than becoming confused with) the icon from view n. Of course, as we scan several times over a scene, we pick up progressively more information, which is integrated into a more complete representation of the scene in consciousness.

It appears that such sensory analyzers and buffers exist for all the major sense modalities—for vision, touch, audition, and so on. Just as we have fleeting visual icons, so do we also have fleeting acoustic and tactile images. These sensory buffers operate in parallel with little interference between them. The interference that arises between monitoring two modalities appears to occur "higher up," in the conscious processor.

2. *Pattern Recognition*

The features extracted from a given stimulus object comprise a sort of coded "description" of the object. In computer jargon, it is a list of attribute–value pairs. In terms of the theory proposed by Anderson and Bower (1973, nicknamed HAM, an acronym for Human Associative Memory), a feature list is a *conjunction of propositions* about the object. For example, the letter E might become coded as "this object consists of 3 horizontal lines which make left-contact with a vertical line." It will be noted here that "feature" has now been extended in meaning to include not only presence or absence of a line but also relationships between the lines and angles and their type of connectivity. Again, work on automatic character recognition (see Duda & Hart, 1973) has provided a variety of mechanical algorithms (in some cases, neural networks) for recognizing such simple relationships among parts.

Following this feature extraction stage, there is an *identification* stage during which the system tries to decide how to classify the stimulus object. In brief, this is assumed to occur by a weighted matching of the current feature list against a likely set of prototypes in long-term memory, with the input being classified according to the name of the best matching prototype (see Reed, 1972). Identification accuracy depends upon the quality or extent of the sensory information extracted in the first stage. Performance also varies with the number of alternative classifications that are expected or are likely given the context. This has been demonstrated many times in degraded perceptual situations in which the size of the expected identification set is varied: accuracy is higher, the fewer the alternative classifications the subject is deciding among (see Garner, 1962). A meaningful stimulus will also be identified more easily if an associated semantic context has been activated near the time the degraded stimulus arrives. Thus, I can more accurately identify a picture of a loaf of bread in a brief glimpse if I am primed to think of kitchen-related items. Or I am more likely to identify the word "America" if I know beforehand that the item fits into the phrase "the United States of _____." Identification accuracy increases directly with the likelihood that one could guess the target item on the basis of the context alone.

Morton (1969) presented a model suggesting how sensory and contextual information might combine in such situations. He supposed that familiar patterns (such as words) have corresponding central units (called *logogens*) accessed directly by the sensory feature lists. Thus, there would be a logogen for common words like EGGS, KEGS, CAKES. By some unknown means, the prevailing semantic context can also supply excitation to particular logogen units. If I have just heard "For breakfast, I had ham and _____" just before I hear a very noisy signal (of "eggs"), the latter signal acting alone might activate logogens that are acoustically similar to EGGS, such as KEGS, LEGS, ACHES. But the context has already placed excitation on a number of logogens related to breakfast foods. The logogens sum their excitation from the two sources, and the dominant one makes its response available to consciousness. Thus, the person will hear the signal as EGGS. The logogens in Morton's theory could be identified with "word nodes" in HAM, the Anderson and Bower theory; that is, a word like EGG would have a central node in the memory network corresponding to a proposition describing how the word is spelled, pronounced, and what it sounds like. What is not specified in either Morton's or HAM's account is the "comprehension machine" that can use a semantic context to activate selected words.

3. *Segmentation and Scene Analysis*

The character identification process described above applies to a single visual object like a printed letter. Having taught the machine to identify single letters, we do not want it to boggle when we show it strings of letters. We want it to treat this pattern as an ensemble of known subparts rather than as an entirely new object. To do so, the pattern recognizer must have the ability to segment or decompose a scene into its several primitive objects, and arrive at a proper identification of these and their interrelationships. Though illustrated for visual scenes, the segmentation problem is just as troublesome for speech recognition since speech tends to flow together in a stream of overlapping phonemes.

The segmentation problem has not been solved in general, but there are several significant programs in artificial intelligence work which appear to be very promising models. To achieve segmentation, a series of passes and interactions are clearly needed between different levels of analysis—with lower-level visual processes that are building up information about elementary lines, angles, joints, etc., and higher-level processes that are testing the developing lower-level descriptions for coherence and plausability. The SEE program of Guzmán (1969) was one of the earlier sophisticated scene analysis programs. It analyzed pictures of three-dimensional blocks scattered about in various orientations and occlusion relations.

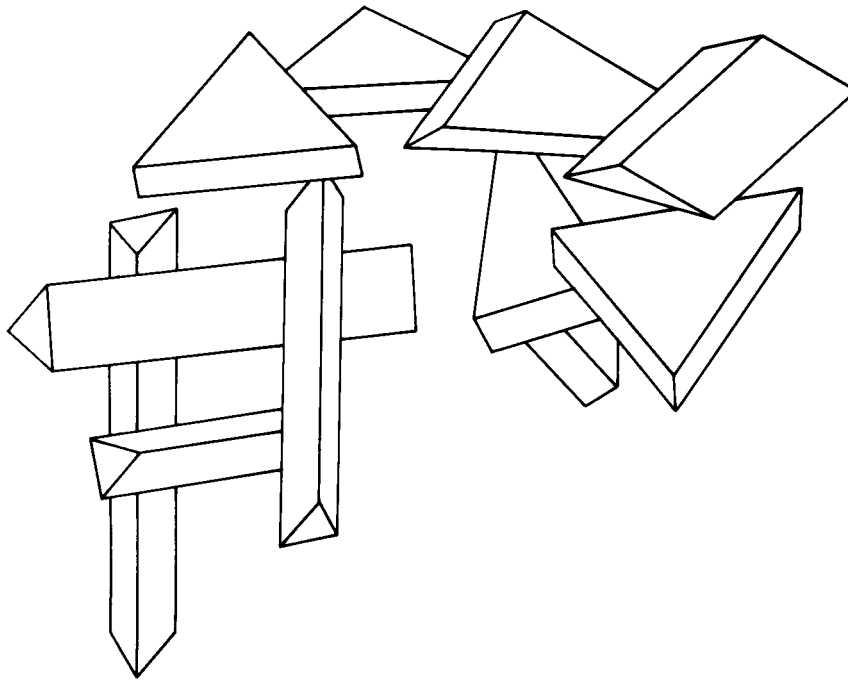

FIGURE 3 A typical toy-block scene analyzed by Guzmán's SEE program. (Reproduced by permission from Guzmán, 1969.)

Figure 3 shows an example scene analyzed by the SEE program. From analyzing a number of such scenes, Guzmán concluded that the most important information arises from intersections, where the surface of one object changes or where the surface of one object is occluded by another surface. Figure 4 shows several of the significant local signs identified by Guzmán, each of which supplies particular types of clues for interpreting the scene. For example, the *L* intersection suggests that the surface to the left (1 in Figure 4a) belongs to a different body from the surface on the right (2). An *arrow* usually suggests two bodies, one with the two surfaces 3 and 4 and the other with surface 5. A pair of complementary *T* joints strongly suggests occlusion of one object (with surfaces 7 and 8) by a foregrounded object (with surface 9). Such local information gathered from vertex clues is combined with more global rules and assumptions about visual scenes in order to arrive at a meaningful segmentation and identification of the parts of the scene.

Winston (1970) has gone beyond Guzmán's program in developing a system that not only segments a visual scene but also describes its objects in terms of geometrical concepts and their interrelations. Thus, Winston's program deals with architectural notions like those in Figure 5. Each of

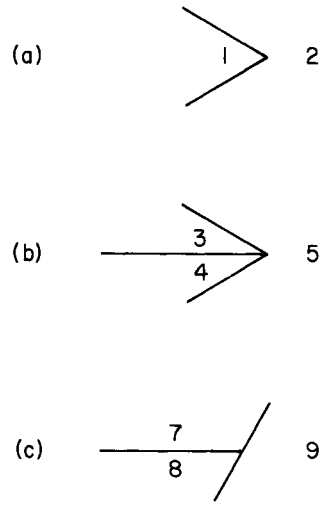

FIGURE 4 Examples of local signs which, in local patterns, signify surfaces, angles, corners, objects, and occlusion relations within block scenes. (a) An L-vertex where two lines meet; (b) an arrow, where three lines meet at a point, with one of the angles bigger than 180°; (c) a T, of three concurrent lines with two of them collinear. (Adapted from Guzmán, 1969.)

these concepts has an elementary structural description. For example, an ARCH is characterized as having three blocks, two upright blocks that are not abutting and are supporting the third. Such structural descriptions (which are labeled relations among elementary concepts) are the output of the perceptual analysis performed by Winston's program. Such descriptions also illustrate the way in which these concepts (of an *arch, tent, house,* etc.) are learned and stored in memory. A test object is classified by matching its structural description to the set of concept descriptors in memory, selecting its name according to the best match (if any).

An important feature of such intermediate-level concepts is that they simplify enormously the description of a scene. For example, consider the scene composed of a tent to the left of a house, with both sitting on top of an arch (see Figure 5). There would be at least seven simple objects—bricks and wedges—in that overall scene, and they would exhibit very many interobject relationships. *Before* the system learns the intermediate-level concepts in Figure 5, all of these parts and myriads of relationships would have to be entered into the description of the scene. However, *after* having learned these concepts, they can be used as units to simplify the overall description to something approximating our English proposition, "a tent to the left of a house both supported by an arch." We can thus see the tremendous economy of internal description provided by our learning of higher-order concepts.

The Winston program is just one of several scene analysis programs available; others exist in conjunction with various "hand–eye" projects and robot projects at several American research centers. Scene analysis is a very difficult problem, and at present only relatively simple scenes are being analyzed. But it is hoped that the methods found successful there will

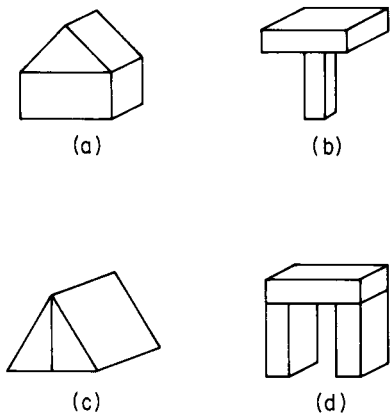

FIGURE 5 Example geometric structures or concepts which Winston's program can recognize and deal with as units: (a) a house, (b) pedestal, (c) tent, (d) arch.

generalize to more complex scenes. Since there is no psychological theory of how people perceive full scenes, these programs are the "best" theories available by virtue of being the *only* theories close to the computing power needed to do the job. Similar progress is being made on speech perception where the segmentation problem is difficult but is being attacked by a "top-down" analysis in terms of semantic constraints (produced by the topic of conversation) on what words are probable in the current context (see Newell *et al.*, 1973; Vicens, 1969).

B. Short-Term Memory

1. General Characteristics

In the course of identifying a stimulus pattern, an internal symbol for it is brought into an active state. Alternatively, we may say that the corresponding symbol has been entered into the subject's short-term memory. Short-term memory (STM) is the active part of the central processor that holds the symbols currently in the focus of attention and conscious processing. Short-term memory need not be viewed as a "place," "register," or "box" physically distinct from long-term memory; STM and long-term memory may only be two different states (or operational characteristics) of the same memory network. For example, the items in the "STM state" may only be that subset of internal symbols (or memory nodes) that have been recently activated. In a computer simulation program, however, entering an item into short-term memory would correspond to placing a token (copy) of its recognized symbol type onto a short list named STM; an alternative representation would be to enter a temporary association from a node named STM to the address of the just recognized symbol type.

The basic characteristics of STM are as follows: (1) as noted, it is the *active* partition of memory; (2) the processor has fast access to the items in STM (that is, retrieval of items from LTM is slower); (3) STM tends

to preserve the temporal order of incoming symbols; (4) STM has a severely limited capacity, about 4–6 symbols or coded items.

Let us comment briefly on these latter three properties. The point about access speed is revealed in experiments showing that someone can respond more quickly to items or facts he has just been thinking about. A likely explanation is this: when a given item has been thought of, its "neural" logogen continues for a while in a state of excitation, with the result that a subsequent input of the item will cause the logogen to pass the excitation threshold needed to fire more rapidly, thus making the item rapidly available to consciousness. This is a plausible mechanism for various sorts of "priming," wherein activation of an item brings it temporarily into readiness to fire again (or "lowers its threshold"). The assumption about access speed in STM is brought out in studies by Sternberg (1969) and others (reviewed elsewhere in these volumes) investigating the time it takes a subject to decide whether a probe or test item belongs to a small predefined set of symbols held in STM. The times increase with the number of symbols in the set, suggesting some kind of a serial scan by the processor over items in STM. While this indicates that the processor does not have immediate access to all items in STM, it is still the case that items in STM are responded to more rapidly than those in long-term memory (e.g., Wescourt & Atkinson, 1974).

That STM preserves the temporal order of arrival of items is clear whenever one is asked to "repeat back" in any order a short string of verbal items he hears or sees. The overwhelming tendency is to preserve temporal order. If one asks for a word series like "eagle, chair, sparrow, soft" to be reported back in an order determined by a semantic classification (for example, furniture, then birds), or another series is to be reported according to rhyming relations, the output time is much slower. It is as though the person has to recirculate, classify, and reshuffle the items slowly to do these latter tasks.

That STM has limited capacity is obvious in terms of the small number of "things" we seem able to keep in mind all at once. Also, our ability to repeat back a random string of items (the "memory span") is severely limited to about seven items. In theory, the span should be longer than the STM capacity since some of that recall is aided by long-term memory (about the test series) build up during initial intake of the series.

2. *Chunking*

The capacity of STM is best stipulated in terms of items or "chunks" of information. G. Miller (1956) first called attention to the fact that verbal memory span is relatively constant, regardless of the "size" of the vocabulary from which the symbol units are being drawn (for example,

letters, digits, syllables, colors, words). But what is a chunk or "unit" for the system? The chunk can only be defined circularly as a stimulus pattern or sequence which the perceptual system recognizes as a familiar, single unit for which there is an internal code (or "symbol"), possibly one designating further semantic structure. Thus, PEN is treated as a unit rather than as three letters, whereas ENP is just three letters. Since JACKPOT has units corresponding to the seven single letters, to four spelling patterns, to two single words, and to one compound word, what determines the "level" of the units that will be entered into STM in this case? Under standard operation, the system normally takes the simplest description of the object (that is, fewest symbols); in this case, that is the internal symbol corresponding to the compound word—the highest level or largest stretch of input over which a successful match occurs. By instructions to the subject, of course, we could bias his "encoding" so that he only pays attention to the letters or the lower-level units.

The units that are identified by the perceptual system depend partly upon the physical characteristics or "gestalt groupings" of the entering elements. The sensory system groups together elements that occur close together in time (or space) and which have similar appearances. Thus, a temporally ordered series of letters with pauses like IBM–FBI–CIO–TV will be remembered much easier than the series IB–MFB–ICI–OTV. Pauses segment the series into four familiar acronyms in the first case but not in the second. But why, really, it this important for memory? It creates a difference in the complexity (or numbers of linkages) of the symbol structures which must be set up in STM to encode and remember the different strings. The well-segmented string eventually is encoded as a string of four tokens of known patterns, whereas the poorly segmented string is represented as twelve tokens (for the individual letters). Therefore, the former string is remembered more easily because it has many fewer connections to be learned.

3. Coding of Information in Short-Term Memory

One of the standard strategies that people have evolved to deal with memorization tasks is to *recode* or translate the recognized symbols in STM from one code domain to a second code domain, hoping thereby to improve their memory for the material. The general model of this coding and decoding is diagrammed in Figure 6. For example, the symbols (or symbol structures) in Domain 1 might be Spanish words that are to be coded into German, or they may be visual letters to be coded into manual sign language, or Morse Code into letters, or printed sheet music into performed music. A common practice seen in verbal learning is for the subject to convert visually presented items into their names. A well-known consequence of this translation is that items then tend to be confused and to

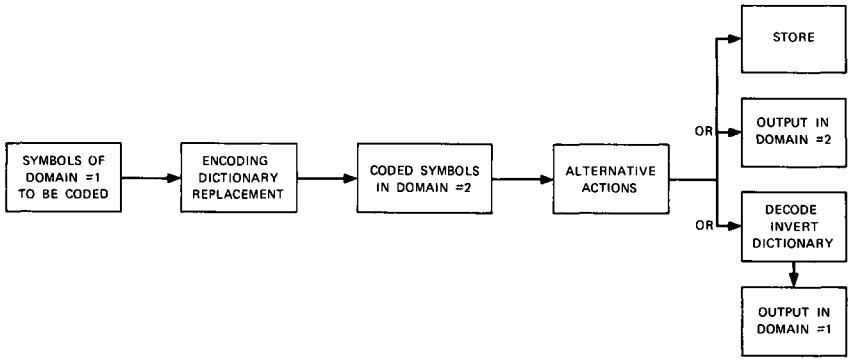

FIGURE 6 Schematic diagram of the encoding process, with the coded symbols then leading to alternative actions depending upon the task demands.

interfere with one another according to their phonetic similarity rather than their visual similarity (Conrad, 1964). If verbalization of the presented items is prevented, then the person seems to rely much more on the visual appearance of the item (see Estes, 1973; Parkinson, Parks, & Kroll, 1971). This suggests that STM can hold symbols (for the order of seconds) that are distinguished by their modality of presentation (for example, auditory or visual).

Different coding schemes may be classified according to whether they preserve, add, or subtract symbols. Some of the mappings are one to one, as indicated above: that is, one visual word converts to one phonetic word, one picture to one name, one sound to one label, and so on. Other mappings are reductive, either because they translate a long string of symbols of Domain 1 into a single symbol in Domain 2 (for example, octal coding, Morse coding, letters to word) or because they select only one distinguishing fragment from the input and discard the remainder. An example of the former kind is when binary sequences like 11001100 are converted into rule-like descriptions of the form "alternating pairs of 1's and 0's (see Glanzer & Clark, 1963). Examples of fragmentary coding would be (1) labeling pictures of different breeds of dogs by the superordinate category, thus discarding distinguishing features, or (2) classifying words according to whether or not they have an initial "bo-" sound, thus aggregating together bow, bowl, boat, etc. Memory for an input may be enhanced by coding to the extent that the code is "shorter" than the original, and permits accurate reconstruction of it. The Glanzer and Clark experiment is a case in point: their subjects were better at recalling those binary series that evoked shorter descriptive codes. In case the code is obtained by ignoring some of the input, then memory for that input will be confused and indeterminate, specifically regarding variations in the ignored parts of the original pattern.

The third class of mappings is elaborative, in which the number of elements appears to be increased during the translation. An example would be coding the name of an event (for example, VE Day) in terms of a set of memories one has associated with that name; another is replacing a word by an image of a known referent and the context of one's knowledge about that (for example, recoding "barber" by imaging the last barber you visited). An equivocal case is so-called "natural language mediators" (NLMs) that adults concoct ro remember meaningless materials like nonsense syllables (see Prytulak, 1971). In order to remember, subjects will add or change letters to convert the syllable into a meaningful word. Thus, LOS converts to LOSe, LOM to LOaM, TAY to sTAY, and so on. Prytulak argued that adults use this coding as a memory aid, since a three-letter symbol structure (L–next–0–next–M) is thus reduced to a meaningful and shorter two-symbol structure (LOAM–without–A). Prytulak supplied a variety of evidence in support of this memory coding, especially for syllables presented at a slow rate.

These several coding schemes generally take time—initial symbol identification, looking up its code in the dictionary, substitution of the code for the original symbol. Consequently, such coding typically occurs only when time is available. But such coding is a trainable skill that improves with practice. Reading—converting the printed to the spoken word or to a semantic meaning—is surely one of the most advanced coding skills we have.

4. Rehearsal Processes in Short-Term Memory

Because most investigations of STM are conducted with adults being exposed to verbal items, there has been much concern in the experimental and theoretical literature with verbal rehearsal (overt or covert) as a subject-controlled strategy for memory enhancement. It is widely agreed that some mechanism akin to subvocal naming of a small series (1) can be consciously initiated, (2) can selectively edit or delete certain elements from study, (3) can maintain a small set of verbal items in STM, and (4) under appropriate conditions can cause information about the rehearsed items to be laid down more permanently into long-term memory.

These assumptions have been summarized most explicitly in the theory of Atkinson and Shiffrin (1968, 1971) and Shiffrin and Atkinson (1969) though anticipated by several others (e.g., Bower, 1964; Broadbent, 1958; Waugh & Norman, 1965). Short-term memory is identified with a rehearsal buffer which is said to have a fixed capacity of r symbols. In a memory task in which the subject is to study and learn a list of unrelated items presented rapidly one at a time (for example, a word list for free recall, or a continuous paired-associate list), the theory supposes that the person

takes successive items into his rehearsal buffer and rehearses them in order to learn them. Each rehearsal cycle is assumed to transfer to LTM some information about the items in STM. Once the arriving stream of items fills up the r slots in the buffer, a new arrival causes a prior member of the buffer to be "bumped out" and replaced by the new arrival. At that moment, the system will stop learning anything more about the bumped-out item; its sole remnant in the system is whatever information may have been built up about it in long-term memory during its time of residence in the rehearsal buffer.

This leads to an "occupancy model" of STM; STM is analogous to a luncheon counter with only r (say, four) seats where customers (items) arrive for servicing, and a set of clocks in long-term memory record how long each customer is allowed to sit before he is bumped off his stool by a new arrival. In a recall task (say, free recall of the list items), it is presumed that an item can be recalled with probability one if it is still in the rehearsal buffer; it it is not in STM, then it may be recalled with a lower probability that is greater the longer its residence time in the buffer (or the greater the amount of information stored in long-term memory). The buffer theorists have never been very explicit about exactly *what* was being stored in long-term memory (they give a functional definition—"it is whatever enables the person to recall better"). In the propositional theory proposed by Anderson and Bower (1973), what is being stored and growing in strength is a proposition of the form, "symbol X appeared with symbols Y and Z in the context of experimental list N." It is then a cluster of such propositions that are accessed when the subject is later cued with the command, "recall all symbols that occurred in the context of list N."

5. *Differential Rehearsal and Priority Setting*

However the model is formulated in detail, it is clear that variations in rehearsal usually lead to differences in learning and performance. It is assumed that selection of a rehearsal strategy and of the set of items to be rehearsed is a plan which is under control of the executive which is trying to maximize achievement of particular goals. Let us briefly mention a few learning effects which apparently can be explained by differential rehearsal.

a. Intentional versus incidental learning. In most measures of recall, intentional learners typically exceed incidental learners. Recording of what they do during exposure to the materials shows that intentional learners are actively rehearsing more. Therefore, motivation to learn and incentives for learning affect what the subject does with the material at the time of

exposure, and his later performance is largely determined by those activities.

b. Selection of a functional cue. In learning with a compound stimulus (for example, a picture of an object and its name), the subject will select one or another element to rehearse depending on how he expects to be tested. If we show him a detailed drawing of a cup, he will look at and rehearse different things if he expects to have to free recall the names of all the objects in the study set than if he expects to have to select exactly that cup from among a set of similar cups (see B. Tversky, 1973).

c. Rewards and punishers as information. We can vary the priority of rehearsal that a subject assigns to two items by telling him at the time of presentation that recalling one will be worth a lot (say, a dollar) whereas recalling the other will be worth very little (a penny). The central processor gives priority to the high-payoff item; it is kept in the rehearsal buffer and newly arriving material tends mainly to displace (and cause forgetting of) the low-payoff items. This analysis is bolstered by the fact that people can remember the approximate payoff they have been promised for remembering a given item. It is likely that the effect of rewards and punishers given after the response in a Thorndikean situation (see the chapter by Nelson, next volume) is largely due to the information they provide about what is to be rehearsed and what can be ignored (see Estes, 1969).

d. Intentional forgetting. Many experiments (reviewed by Bjork, 1972, and Epstein, 1972) have shown that subjects can selectively "delete" items from their current STM. This can be done simply by telling the learning subject soon after he has entered one or more items into STM that he will no longer be held responsible for remembering those items (that is, he can "forget" them). This forget instruction has several effects: it causes the subject to cease rehearsing the forget items and to devote more rehearsal to the items on which he expects to be tested; the forget signal is also a stimulus event that can become associated with any items with which it was paired, so that the person may be able later to indicate whether a given item was a "forget" or a "remember" item; and earlier "forget" items may be rejected as possible associative contacts to help organize later "remember" items in a free recall list. The literature cited may be consulted to become acquainted with the experimental data surrounding investigations of intentional forgetting.

e. Priority of distinctive items. A result by von Restorff (1933) and many others is that an item that is distinctively different or surprising in relation to the other items in a learning series will tend to be learned first, but it does so at the expense of neighboring items of its series. An experiment by Ellis, Detterman, Runcie, McCarver, and Craig (1971) illustrates

the point: subjects saw many series of slides of simple objects, recalling their names after each series. For the critical series a shot of a pornographic nude was shown without comment in the middle of the picture series. Plotting serial position curves for free recall, the nude picture was recalled 100% (well above the corresponding picture in the control list), whereas pictures adjacent to the nude picture were recalled worse than their corresponding controls. An explanation in terms of the buffer model is that the startling item caused rehearsal of just-prior items to cease, whereas subsequent items were either not entered into STM or were assigned low-priority rehearsal scores because the "nude picture" code remained in a state of high activation for awhile.

f. Instructional differences in rehearsal. The experimenter can instruct subjects to rehearse the series of study items according to distinctively different patterns, for example, one at a time, by nonoverlapping triplets, by cumulative addition of the next (see Atkinson & Shiffrin, 1971). These instructed variations in rehearsal produce differences in later free recall as predicted by the buffer model (for example, reduced primacy effect in "one at a time" rehearsal, enhanced primacy in cumulative rehearsal, and triplet-clustering in the third condition).

g. Individual differences in rehearsal. There are large differences among individuals in memory. Young children, by and large, remember less well than older ones; children with high IQs usually remember more than those with low IQs—which is why most IQ tests include memory subtests. There are doubtless a variety of causes for these differences. One difference, for example, is that the older, more educated person may have become familiar with more subsequences (chunks) of the test materials, so that he can recode some lengthy series into a slightly shorter sequence of internal symbols (for example, letters to spelling patterns). Another likely contributor to the difference in memory is the fact that older, smarter children use more efficient mnemonic methods. Flavell and his co-workers (e.g., Flavell, Friedrichs, & Hoyt, 1970) have found marked differences among children in their spontaneous use of verbal naming and sequential rehearsal of object series they were to remember; older children tend to rehearse more and remember more. If younger children are induced to rehearse more efficiently as they study the series, then their later recall equals that of the older children. In other work, it has been found that nursery school children do not behave differently under instructions to just "look at" versus "memorize" a picture series, whereas older children employ deliberate memorizing strategies when given the latter instructions. This suggests that memorization strategies develop gradually as a way to deal with a future contingency, one that small children do not yet understand (see Appel, Cooper, McCarrell, Sims-Knight, Yussen, & Flavell,

1972). Similarly, Belmont and Butterfield (1971) found marked differences in the way normal children and retarded children pace themselves through a self-presented list of items they are to memorize. Normals often do "grouping" and "cumulative rehearsal," returning repeatedly to the beginning to go over the entire series up to the present point as though testing themselves. In contrast, retarded children tend to zip rapidly through the study series without substantial pauses. There are correspondingly large differences in the two groups' recall of the series, particularly in the early ("primacy") items of the series. If the retarded children are trained to use the cumulative rehearsal strategy, with longer pauses for self-testing, their performance improves to become much more like that of normals on this task. This research area, comparing children of different mental ages on their mnemonic methods has just begun to be explored in depth over the past few years. The practical hope, of course, is to find teachable methods that will accelerate memory development in small children and retarded children.

6. *Shortcomings of Verbal Rehearsal Models*

The foregoing is just a partial list of the sorts of results that are explicable by the notion of a verbal rehearsal buffer that is under strategic control. For other examples, see Atkinson and Shiffrin (1968) and especially Rundus (1971), who devised the method of making rehearsal overt in order to examine whether the mnemonic influence of several variables (such as the spacing between presentations of a repeated item) was mediated through their direct effect on amount of rehearsal of the item in question. These methods have indeed supplied us with very valuable information.

However, there are several shortcomings of the supposition that short-term memory is coextensive with the verbal rehearsal buffer as described in these models. A first difficulty arises when we realize that we have millions of memories for nonverbal events or items which we have not verbally described—for paintings, music, scenery, all varieties of sensory magnitudes such as the pitch and loudness of tones, the length and grayness of lines, pressures, smells, tastes, etc. Of course, some of these we do classify verbally (for example, "it was about a three-inch line"); such a verbal description clearly can be rehearsed, and it will influence our memory (or reproduction) of the stimulus. But it is just as obvious that we can remember something about nonverbal events even if we do not (or can not) describe them. They seem to be represented in terms of analogical structures that are not necessarily connected to the verbal production system.

It is a moot question which modalities have a usable covert rehearsal mechanism. It would appear that some kind of motor reproduction system

is required in order for covert rehearsal to suceed in maintaining a stimulus sequence in STM and to aid its transfer to long-term memory. For example, musical melodies are doubtless stored in an abstract (relational) symbol structure bearing no relation to the verbal system; but we do have a motor output system—humming or silent whistling—that allows us to regenerate a once-heard melody, and presumably thereby store it in memory better. But is there a similar covert rehearsal mechanism for odor sequences, for dance steps, or for sequences of tastes and touch patterns?

A second problem with the rehearsal buffer notion is that it assigns far too much weight to "dull, dead repetition" of the items themselves as the sine qua non for entry of items into more permanent memory. But subvocal naming of the items may only be the observable tip of an "iceberg" of deeper semantic processing. Specifically it now looks as if the transfer of information about items to LTM requires that those new items be related to familiar things in LTM and that interrelations among the new items be noticed and used for learning.

This viewpoint has been suggested by several considerations (see Craik, 1973; Jacoby, 1973; Mandler, 1967). First are increasing reports of situations in which verbal rehearsal of materials has virtually no influence on long-term retention of that material. Craik (1973) has distinguished Type I (superficial maintenance) rehearsal from Type II (semantic, organizational) rehearsal. Type I rehearsal causes verbal items to be maintained for a brief time in phonetic recirculation, but it results in very little storage of those items in LTM. Type II rehearsal involves more activation and consideration of the semantic meanings and associations of the items, trying to interrelate several of the to-be-learned items together in a meaningful sentence, image, or the like. An elementary experiment by Jacoby (1973) will illustrate the distinction. Adults listened to five unrelated words (well below their memory span) and either recalled them immediately or spent 15 sec rehearsing them subvocally before recalling them. Though recall was perfect in either case, items receiving the 15 sec of rehearsal certainly had a much higher frequency and duration of residence in the STM. After doing a number of such short lists, Jacoby surprised his subjects by asking them at the end of the session to recall any and all the items they could from the entire experiment. Significantly, recall of the multiply rehearsed items was no better than recall of the immediately recalled items—a clear failure of effect of repetition. On the other hand, we know that if, during that 15-sec interval, the subject had been asked to organize the five list words into a story or into an interactive scene in imagery, then later recall would have been greatly enhanced. A number of such demonstrations have shown us that people have a way of temporarily holding information in a phonetic form that produces almost no longer-term memory for the material. The distinguishing mark of experiments illustrating this point is that

the subject is not expecting the long-term retention test; he is, in fact, led to believe that once the material is "dumped from STM," then it can be forgotten. So it would appear that maintenance of items in STM is not a *sufficient* condition for the transfer of much information about them to LTS.

The converse side of the coin is that "residence time" of an item in STM is not so strong a predictor of long-term retention as is the type of processing the item receives during its residence in STM. Again, recent experiments by Craik (1973) and Jacoby (1973) prove this point as did earlier demonstrations by Mandler (1967), Tulving (1966), and Bower and Winzenz (1970). In Craik's experiment, the subject saw a word briefly after being set to judge it either in terms of its physical appearance (upper- or lower-case letters), its sound (rhymes with *hat*), its semantic category (is an "animal"), or its fittingness in a sentence frame ("blind as a _____"). Later free recall and recognition indicated much better memory for words that had been processed to the "deeper" semantic levels. An explanation for these results is currently being pursued. A likely explanation for the recognition memory results would begin by noting that recognition memory depends upon retrieval by the test item of information about the context of study of that item. We may suppose that the test word is better able to produce recall of its study question following the semantic-classification task simply because the word-to-category link has been primed by the study question and it preexists in memory, is strong, and is unique, whereas the word-to-type-font link (for the "shallow" processing) is novel, weak, and interfered with by many other items (for example, many items appear in upper case letters). Similar factors would imply better free recall for semantic than for case-queried items if it is supposed that semantic categories serve as the normal retrieval cues for free recall (and these *were* unique cues in Craik's experiment), whereas the type font or the sound of the item is not a self-generated retrieval cue. If this line of explanation were to be successfully extended, then a plausible interpretation of the "depth of processing" result is that it reflects the retrievability of the "question-to-item" event, which in turn depends upon the way the question biases the encoding of the item and the quality (uniqueness, prior associative connection) of the question-to-item association. These are currently active items on the research agenda.

C. Working Memory

To continue with the system architecture of Figure 2, we use the term working memory (or intermediate-term memory, ITM) to refer to the memory structures which maintain information about the local context, but information which is neither in the focus of active memory nor in the

distant edges of long-term memory. One of the primary functions of working memory is to build up and maintain an internal model of the immediate environment and what has been happening in our world over the past minute or two. We may think of the working memory as containing a description of the *setting,* framework, or context within which the more dynamic alternations of the world before us are taking place.

Several arguments indicate that we must have these temporary "world frames." First, visual perception is itself fleeting, skitterish, darting hither and yonder about a scene. But yet we do not every 200 msec recognize and construct the entire scene anew. Rather we integrate the information from those glances into one sustained image or model of the scene "out there" and our place in it. The model tells us the objects out there and their distribution in space with respect to ourselves.

This local model serves as a framework within which dynamic (small) changes are recorded—old objects are deleted, new ones are added, things change their relative positions, etc. But only the new, altered information must be focused on (in active STM) and entered as a new symbol structure or proposition into working memory. That is, new information "updates" rather than completely casting aside the current model.

These temporary models serve as a context for perception. In doing so, they constantly call upon information stored in long-term memory concerning the nature of objects and spatial relationships. For example, if I have just walked past a man who is climbing into his car, I am not startled to hear behind me the noise of his motor starting up a few seconds later. We "perceive" the rumble as a car motor rather than thunder or a building collapsing because our contextual model makes "motor noises" plausible. It is on the basis of long-term memory that I know that the lamp shade suspended apparently in midair across the room from me is in fact supported by that steel rod disappearing beneath it, and that a wire from the rod to the wall plug carries electric current to the bulb. So I am prepared to see certain things if I go over there and look under the lamp shade; also I expect certain changes in my "world model" if I pull out the electric plug or cut the wire. The point of such elementary examples is to say that perception is very much a matter of "predicting" the results of my interactions with my environment, and it is my current world model (based on partial sensory information) that enables me to bring the appropriate knowledge to bear on this issue.

Another phenomenon illustrating the need for an intermediate-term memory is in keeping track of what topics and referents have been recently mentioned during linguistic discourse. This is required in order to find the appropriate referents of anaphoric and pronominal forms. Also, it is used in finding the correct sense of ambiguous expressions. If I mention that

"John is a baker," later in the conversation I can refer to him as "the baker" or simply as "he,"; but to do so requires the listener (and me) to have a memory that keeps track of who is under consideration. Similarly, if the discourse is about a birthday party for little Billy, we must have that theme carried along in working memory to activate belief structures which provide proper interpretations and reasons for ambiguous sentences like "I'll buy him a top" or pronoun reference as in, "Don't give him one. He already has a top. If you do, he'll make you take it back." In this last example, the problem is to figure out that the *it* does not refer to the top that Billy already has but to the hypothetical "one" you are considering giving to him (see Charniak, 1972, for many examples). To summarize the point here, a working memory is necessary to keep track of foregrounded contexts in dialogues.

Another plausible use of a working memory is for it to hold proposed deviations from our actual world model that arise as we are thinking and trying to *imagine* the consequences of making certain alterations in our local environment. For instance, while looking at a configuration of chessmen, the master will try out (in imagination) a sequence of moves, entering the alterations in board positions of the relevant chessman in his temporary working model of the board, and then evaluating these projected imaginary states to decide upon a move. Similarly, an interior decorator rearranging furniture in a living room can use his working memory as a temporary imaginal model ("sketch pad") for trying out and evaluating various arrangements.

Finally, laboratory psychologists might identify the contents of the subject's working memory with the full set of propositions he knows about the experimental context in which he finds himself while engaged in, say, a nonsense-syllable learning experiment. At any moment we could ask him, "Where are you and what have you been doing?," and (with eyes closed) he could relate his complete spatial orientation to us, the room layout, where the materials-to-be-learned are displayed, what types they are, etc. He could also tell us what his instructions are, his strategies for dealing with the learning task, and how he evaluates the task as well as his performance. All of this would serve as an immediate mnemonic background or context within which his active short-term memory has just recorded episodic propositions like, "most recent letters in the memory-drum window were X–then–P–then–H–then–R." Through instructions, the subject knows that when we ask for recall, he is in fact to enter the retrieval cue, "most recent letters in the memory-drum window were ???." To use our earlier terminology, the working memory contains (among other things) the *plans* that the subject has initiated for dealing with the task we have set for him.

D. Long-Term Memory

1. General Considerations

The long-term memory (LTM) is assumed to be the "repository" of our more permanent knowledge and skills. It essentially includes all things that are in memory that are not currently being used. The information structures in LTM includes such general classes as the following:

1. Our spatial model of the world surrounding us—symbol structures corresponding to images of our house, city, country, and planet, and information about where significant objects are located in that cognitive map.
2. Our knowledge of physical laws, cosmology, of the properties of objects and things.
3. Our beliefs about people, about ourselves, about how to behave in various social situations. Our values and the social goals that we seek.
4. Our motor skills for driving, bicycling, shooting pool, etc. Our problem-solving skills for various domains. Our plans for how to achieve various things.
5. Our perceptual skills in understanding language or interpreting paintings or music.

Even at such a general level, the listing is hardly complete. Some of the information structures, when activated, produce a conscious experience of sensory imagery, some produce only a verbal flow (words of a memorized poem); others produce no introspective feedback whatsoever (for example, the "rules" we appear to follow in understanding speech).

At present there are just the beginning of a few models of LTM being proposed, specifically those by Anderson and Bower (1973), Kintsch (1974), and Norman and Rumelhart (1975). To these may be added the efforts of many computational linguistis (e.g., Quillian, Schank, Winograd) who have worked on models for language understanding; such models require a highly organized memory, one that interacts continuously with a syntax parser in arriving at an internal description of an incoming sentence. The latter have been concerned with constructing memories suited for language parsing, but have not addressed themselves specifically to issues of learning, retrieval, forgetting, and such matters. All the systems are, in fact, oriented toward dealing with *sentence* inputs—with propositional "stimuli"—and they discuss other matters only by metaphorical extension of the way their systems deal with English sentences. All of these theories assume that LTM is a conceptual network, which serves as the" data base"

into which information (propositions) is inserted through the working memory and from which the central processor retrieves answers to questions. Conceptual networks were hit upon as a way to implement some general properties of human memory; therefore, to explicate such network theories, let us consider some of these properties.

2. *Some General Properties of Human Memory*

We list a few characteristics of the memory required and then describe a few functions that we would want a memory model to perform.

a. Associative retrieval. A first property is that information in memory should be retrieved in response to information that is associated with the input stimulus. To use a metaphor, the location of information in memory is specified in part by the contents of the item itself. Thus, a given symbol description (e.g., GLURK) will gain access to a spot in LTM at which would be collected whatever information (if any) that the person had associated to that stimulus pattern. The EPAM model of Feigenbaum (1963) is one realization of a content-addressable system. (In EPAM, the memory location corresponding to a stimulus item is arrived at by sorting the description of the stimulus through a series of feature tests.) Computer memories are not organized in this manner: rather data is placed seriatum in cells having locational addresses, but the address is unrelated to the datum stored there. An associative or "content-addressing" system can be approximated on current computers by "hash coding"; in such schemes the memory address containing information about an item is computed as some "hash function" of the item's contents.

The important issue for models of human memory concerns the units level at which the system is content addressable. Models like HAM assume direct access to sensory features and sensory relations; but given a perceptual description of a stimulus object, the system makes use of associative retrieval from there onwards. That is, a sensory feature pattern might "associatively retrieve" a word, and then a compound semantic idea (for example, horse meat, golf ball).

b. Concepts as nodes. The basic elements of memory are concepts (symbols) and relations between concepts (symbol structures). A concept may be a perceptual primitive (for example, vertical line, blue, horizontal movement), an actional primitive (for example, move hand, grasp, chew, suck, release), a primitive relation (for example, inside, beside, part of, has as parts, is a member of, above), or a higher-order concept built up by relations among these primitives. The concepts can stand for both generic terms (cups, pencils, governments) as well as individual constants (specific people and places). The "meaning" of a concept is given partly by

the configuration of its relationships to other concepts and partly by the referential conditions necessary for the proper use and application of the term. In our representation and in diagrams, concepts are represented by nodes (or cells in a computer memory) and relations between concepts as arcs, arrows, or labeled associations between the nodes.

c. Information as new conceptual configurations. In such a conceptual network, the learning of a new fact or new concept is solely a matter of recording its representation, which is in terms of a specified configuration of relations among already known concepts. Thus, for example, the concept of a helicopter would be introduced by listing its general type (airplane) and function, enumerating its distinguishing features (for example, propeller on top, capability for vertical takeoffs and landings), perhaps storing something equivalent to an "image" of the object. A new name may be associated to this concept, but the name is itself just another symbol structure (constructed from letters or phonemes). Most of our concepts do not have proper names but only definite descriptions (for example, "the person who stole my bicycle"), but they nonetheless would be represented in memory as a single symbol designating an entire symbol structure.

d. Concatenation principles. Our memories must have the capability to build up arbitrarily complex concepts from simpler concepts by use of simple rules of combination and concatenation. To do this, we must be able to say something further about any concept—that is, to attach another "relation-plus-concept" to any arbitrary concept. If I have the concept of the dog instance, Fido, I can predicate of him that he bit Harry, that I dislike him, that Jane owns him, and that Jane dislikes the fact that I dislike Fido. In these latter cases, notice that a *proposition* is itself being taken as an object (a "concept"). By representing propositions as nodes in memory, we are able to predicate about them as if they were simple concepts. From the viewpoint of the central processor, this means that a known proposition can be represented in working memory by a single token of its corresponding node in LTM. For example, we can deny a proposition by predicating "is false" of the proposition node.

e. Creating tokens from types. Any given concept, feature, or attribute will be used in many contexts, in encoding a great many events. For example, every time we store a proposition describing an event involving a new instance of a dog, we want to be able to create a new concept (the dog-instance) and say something about the event in which he occurs. We cannot predicate this particular event about the class of dogs; rather we must create in memory a *token* or *new instance* of the class of dogs, and attach the predication to this new instance. In the Anderson and Bower (1973) theory, this type–token link is simply one of set membership: Node

3714 (say) in memory is created to stand for a new member of the class of dogs, and we can then predicate many things about this dog instance (for example, its appearance and episodes in which it occurs). This capability for creating tokens of preexisting types is in fact the same capability noted earlier, of being able to create new concepts (for example, dog-3714) in terms of their relations to old concepts (for example, "dog–3714" is a dog).

f. Types of conceptualizations. The current models like HAM assume that the input to LTM is in a form equivalent to a proposition or a conceptualization. That is, LTM is capable of recording only inputs which have been "propositionalized" in an appropriate manner by the perceptual system. But this allows fantastically wide latitude. In the Anderson and Bower theory, a proposition is a subject–predicate construction put together according to certain rules. But such stipulations only invite one to supply the semantics of permissable combinations. Among the obvious types of semantic combinations which qualify as conceptualizations would be the following (suggested by Schank, 1973a):

ACTORS perform ACTIONS
ACTIONS have OBJECTS
ACTIONS have INSTRUMENTS
ACTIONS have RECIPIENTS
ACTIONS may have DIRECTIONS
OBJECTS can relate to other OBJECTS
  these relations are: POSSESSION
  LOCATION
  CONTAINMENT
OBJECTS can have ATTRIBUTES
ACTIONS can have ATTRIBUTES
ATTRIBUTES have VALUES
CONCEPTUALIZATIONS can have TIMES
CONCEPTUALIZATIONS can have LOCATIONS
CONCEPTUALIZATIONS can CAUSE OBJECT'S ATTRIBUTES TO CHANGE VALUE
CONCEPTUALIZATIONS can ENABLE other CONCEPTUALIZATIONS to occur
CONCEPTUALIZATIONS can serve as REASONS for other CONCEPTUALIZATIONS [p. 5].

These presumably serve as some of the "syntax rules" for composing allowable conceptualizations. During conceptual development, particular concepts come to be classified as actors or objects, and particular actions come to be classified according to what conceptual cases they require (for example, "*physically give*" requires an actor, object, recipient, and direction of transfer of possession). Schank proposes that starting from a few primitive semantic notions and such rules as above, the infant acquires through perception a fragmentary conceptual knowledge of its environment onto which the structures of spoken language are then grafted. Then the two systems of language and world knowledge bootstrap each other up to the sophisticated cognitive machine that we are.

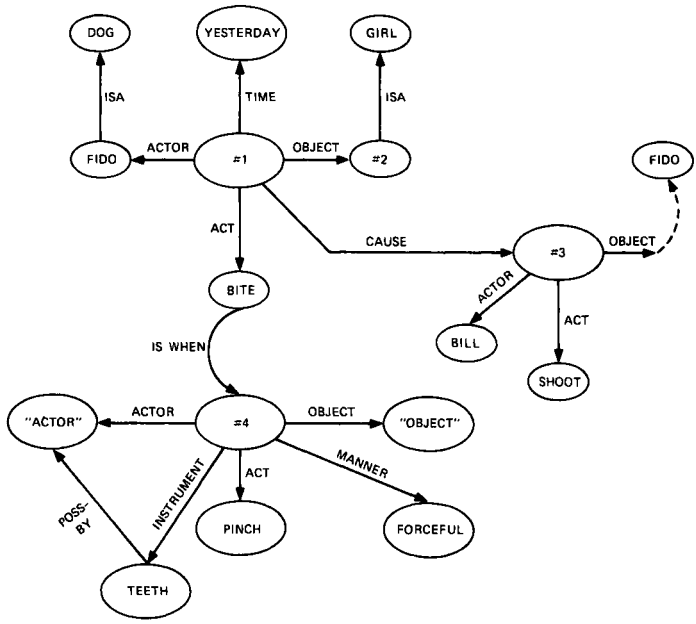

FIGURE 7 Diagram of the associative structure set up in memory to encode the conceptual content of "Fido bit a girl yesterday, so Bill shot him." The arrows are labeled associations (relations). The definition of *bite* is also shown. Definitions and knowledge about each concept is not shown to reduce clutter.

Putting aside the validity of such conjectures (or indeed, the more basic question of how to test them), it is fair to say that the memory network theorists seek a set of basic concepts and combination rules which will allow their system to represent practically any statement, event, or state of affairs. In particular, happenings or *events* in the world are to be represented in terms of a new configuration of relationships between tokens of preexisting concepts. To illustrate, look at Figure 7, which is a diagram representing the event "Fido bit a girl yesterday, so Bill shot him." Concepts are represented by circles, and relations by arrows. Node 1 represents the first event of Fido biting a girl-instance yesterday, and this event is said to cause the event in Node 3, that Bill shoots Fido. The small piece of network beneath *bite* illustrates the type of lexical decomposition into meaning components that linguists propose for most verbs. Thus, *bite* means for the actor to forcefully pinch the object using his teeth.

Such representations allow for retrieval of information by questions (probes) which stipulate part of the memory structure. Questions act like compound "retrieval cues," to use terms familiar in the memory field. If we ask, "Who did Fido bite?," the entry points (Fido, bite) retrieve Node 1 from memory; the interpreter then asks for "OBJECT-link of Node 1," which returns with Node 2; Node 2 is then found to be a girl instance—and

that determines the answer, "Fido bit a girl." If we ask "*Why* did Bill shoot Fido?," the system looks up the conceptualization of the query, getting a match to Node 3; it then backs up the causal link to Proposition 1, which permits it to construct the answer, "because Fido bit a girl yesterday."

g. Flexible retrieval. Another property we require of a humanoid memory is that a given memory be accessible by means of a variety of different cues. The supposed advantage of encoding events into *conceptual* representations rather than literal word strings or pictures is that the issue of flexibility of input or output by paraphrasing is handled in the input parser or the output generator (which are peripheral to the memory itself). Thus, I might have *seen* Fido bite the girl, and later learn verbally that because of that event Bill shot him; but my core memory representation is supposedly about the same as above. Similarly I would get the same memory structure if I had heard "Fido was shot by Bill because he found out that Fido had bitten a girl yesterday," or some other such paraphrase of the atomic propositions. Conceptual networks also have the potential flexibility for retrieval by paraphrases. Thus, I could ask, "Did a dog bite a girl?" requiring simple looking up of set membership for Fido; or I can ask, "Who forcefully pinched whom yesterday with his teeth?" and have hopes of perhaps getting a sensible matching answer at the level of concepts.

h. Ability to evaluate inputs. Another basic trait of human memory is that we constantly evaluate our perceptual or linguistic inputs to check for their consistency with what we know. We evaluate the truth of what is told to us; we answer "true" or "un-hm" if an input statement matches its conceptual counterpart in memory, and answer "false" or simply balk if we come across contradictory information ("Fido bit a boy yesterday"). Similarly, our memory provides an almost immediate decision that an input violates syntactic rules or semantic selectional restrictions ("Fido bit a ritual yesterday"). The exact means by which these contradictions and plausibility judgments are arrived at is a topic of continuing concern in research on semantic question answering.

i. Inference potential. It has long been recognized that many of the questions humans answer (read "much of the information retrieved from memory") is based on inferences from what we know rather than simply being a direct reading out of stored propositions. One way to view this is to think of the facts stored in our memory as like "axioms" of a logical system, so that questions are answered (if not directly matched) by applying a set of "subjective derivation rules" to our knowledge to come up with the answer. Thus, we derive that ostrichs have lungs because they

live above ground and all such animals have lungs; or derive that Spinoza had an elbow because he was a man and all men have elbows. Such examples use the logical properties of relations like "set membership" and "has as parts"; that is, if x is a member of set y, and all y's have property p, then x has property p. A large number of relations (for example, "is located at") have some such logical properties.

A current focus of theoretical and experimental activity concerns how such inferences are made in semantic memory, particularly for set-membership statements ("a crocodile is a reptile") and for simple "has property" statements ("a crocodile has a stomach"). The competing theories must take account of the quantifier of the statement ("all, many, some, few, or no trees have leaves"), whether or not the subject and predicate are closely related in meaning, and, of course, whether the statement is true or false (for a good example, see Glass & Holyoak, in press; Glass, Holyoak, & O'Dell, 1974). Theories differ according to their assumptions regarding the organization of our conceptual knowledge and the search mechanism that operates over that network. Though the best model for the domain is in dispute, progress is clearly being made in research on the topic.

The discussion above refers to inferences drawn from long-term semantic memory. But another salient feature of humans is that they use a current event, episode, or statement to set off a chain of inferences about the context and situation described by the statement—inferences about the probable antecedents or reasons for the event, the probable consequences of it, and so on. We fit the event into our "model of the world," and are then able to expand it by inferences about various consequences. Consider just the unfortunate Fido incident above. We may infer any and all of the following:

1. Fido has teeth, probably sharp, and he has jaws.
2. His teeth made physical contact with some body part of a girl.
3. This probably caused the girl to feel pain and to cry.
4. Bill learned of this event.
5. It made Bill angry at Fido.
6. He decided to dispose of Fido.
7. He came in possession of a gun.
8. The gun was loaded.
9. Bill aimed it at Fido and pulled the trigger.
10. A high-velocity bullet passed through Fido's body.
11. This probably caused Fido to become dead.

This is but a brief listing of the myriad relations an average listener can "derive" or infer from the sentence said several pages back. The derivations are obviously of several different kinds: some come just from the meaning

of the action verb, some from the preconditions necessary for the given action to occur (for example, Fido and the girl had to be in the same location); other inferences come from our layman's postulates about emotional or causal consequences of actions and our reasons for actions. Abelson (1973) has written cogently on these issues, and Rieger (1974) has developed a simulation program that generates a large network of such inferences surrounding a given input sentence. Rieger argues that such inferences are frequently needed to follow the reasoning that weaves itself throughout even the most simple conversation. For example, if I know that Bill does not own a gun, then when I hear the original sentence the inferences stumble upon a "gap," which leads me to ask something like, "How did Bill get a gun?"

A point glossed over in the above illustration is that we have introduced "cause" as a relation between two propositions—between two events or states. Other propositional connectives would be *and, then, while, or,* etc. From such connectives one can represent stretches of coherent text much longer than single sentences. Thus, one can imagine giving a hierarchical description to stories. Speaking abstractly, a given story might consist of, say, Episode 1, then Episode 2, which cause Episode 3. But Episode 1 in turn consists of Event 1 together with Event 2, both causing Event 3, which completes Episode 1, and so initiates Episode 2 of the story, etc. Analysis of the conceptual structure of texts and stories, and their encoding into memory, is just beginning (see Crothers, 1972; Frederiksen, 1972; Meyer, 1974; Rumelhart, 1973), but this is clearly going to be a very important area for future research.

A current topic of interest to linguists is the issue of how to classify and decompose verbs into a small set of primary meaning components. Schank (1972, 1973b) and Norman and Rumelhart (1975) have devoted major efforts to this topic. The motivating philosophy is much like that behind early chemists' search for basic atoms from which complex molecules would be composed. One implication, for example, is that if different meaning components become available at different stages during cognitive development, then in learning his language a child should progress through an orderly set of "stages" at which he understands progressively more complex verbs (for example, *have,* then *give,* then *sell*) while being confused among verbs that use semantic distinctions he has not yet acquired.

3. *Learning in Conceptual Networks*

The treatment of cognitive learning in these network models is perhaps obvious given the foregoing background. When a new fact (proposition, statement, scene) is entered into working memory, we suppose that it corresponds to a structured conceptual tree diagram such as shown in Figure

7. What is "in working memory" at this time are the tokens of the LTM concepts and the temporary connections (labeled links, associations) that are to be established in order to remember the fact. Learning consists in recording these linkages into long-term memory.

An initial question is, "When will any learning be initiated?" This would seem to depend on the "informativeness" (or unpredictability) of the stimulus sequence or event under consideration. We assumed earlier that each incoming proposition is first checked for whether all or part of it is already known (is stored in LTM). If it is a familiar fact, then little is stored about it—perhaps only the further fact that it was mentioned in the present context. Since working memory is constantly using part of the current world model and LTM to predict forth-coming events (for example, if lightning, expect thunder), the foregoing assumptions indicate that the occurrence of predicted events is easily perceived but unlikely to activate any efforts toward new learning of coincidental events.

If an incomng statement finds a partial match between its information and an existing tree structure in LTM, then that matching structure is used as the node onto which the difference information ("new predications") are attached. If I already have stored that Fido bit a girl and I am later told that the girl Fido bit was little, rich, and blonde, I do not rerecord all the nodes and links encoding the original fact. I simply add to that structure (specifically, to Node 2 in Figure 7) the new attributes asserted about the girl instance. The important assumption here, brought out particularly in the learning routines of the Anderson–Bower theory, is that, whenever possible, incoming information is recoded and recorded by reusing and modifying parts of already known information structures. This is economical in eliminating redundancies of a certain sort. It also illustrates why highly associated material is so easily learned (for example, compare the learning of "the doctor cured the patient" versus "the doctor slapped the hostage")—because most or all of it is already known. The hypothesis also accommodates our tendency to *assimilate* new information in terms of familiar conventions and stereotypes (see Bartlett, 1932). We may store an event as "Event 7 has familiar Description 172 except for Differences 1, 3, 6." But then we forget the differences (they were not successfully recorded in LTM) and so reconstruct our memory of the event or story in familiar, conventional form.

Let us consider finally the case where the incoming statement expresses totally new propositions—they are sensible but not already known in LTM. How then does learning proceed? Presumably by rehearsal of the network of relations among the concepts and by thinking about and perhaps "imaging" the referents of the concepts and their stipulated interrelations. As a *quantitative* description, it may be assumed that in each small unit of time as the person is concentrating on the proposition, there is a fixed

probability that any given link in the input tree will be recorded into long-term memory. Given a fixed study time per sentence (typically, 3–10 sec in most learning experiments), the the number of relational links that are durably formed in a tree is a probabilistic matter. That is, the full statement is not learned all or none as a unit; rather some subparts may be learned before others (see Anderson & Bower, 1973, Chapter 10).

The result of this stochastic learning assumption is that there will be many "fragmentary trees" in memory that result in partial, vague memories. And such fragmentary memories do indeed exist, either when we introspect or examine experimental recall results. Thus, you might remember that "Fido bit someone, for which he was shot," but forget who was bitten, when it occurred, and who shot him. Such fragments, plus inferential elaborations done at the time of learning or recall, provide the material for plausible, imaginative reconstructions done according to normative knowledge. For example, I might elaborate my fragments to "Fido probably had rabies, was running wild in the park where he bit an old man, and so was shot by a passing policeman who witnessed the attack." Our conventional knowledge and stereotypes have elbow room to intrude in just such cases.

The propositional learning model given above is hardly profound, but there is empirical evidence in its favor (see Chapter 10 and 11 of Anderson & Bower, 1973) and no intuitively compelling alternative has been proposed. A problem with it is that it seems testable only given the naive assumption that a word given alone as a retrieval cue (for example, *bite*) is interpreted in the same semantic sense as it was when it occurred in the context of the study sentence (as in "dog bite"). In fact, however, subjects may occasionally think of a different sense of *bite* than that encoded during input (for example, *bite* can also mean to beg from someone, or to sting, or to clip one's speech). But these are second-order deviations from the basic model of probabilistic link formation of learning (see R. Anderson & Ortony, in press, for counterarguments).

E. Decision Making

An organism is faced with a "decision" when he has two or more alternative courses of action for each of which he must evaluate and compare the expected consequences. In a simple experimental situation involving two stimulus–action–outcome alternatives, the person obviously tries to accumulate knowledge allowing him to predict the payoffs or penalties following each stimulus—action combination. He essentially learns such correlations by accumulating rehearsal on propositions such as "pressing the left key to the cue MIB results in five points." For a rational decision, this bit of knowledge would be needed along with learning of a complementary proposition such as "pressing the right key to the cue MIB results

in three points." Given these two bits of knowledge and a higher valuation on five than on three points, then presentation of MIB will retrieve the two propositions and the rule of "choose the more preferred outcome" will cause the central processor to issue the command "execute a left-key press." Such is a simple matter.

More complex decisions involve either choices between two large commodity bundles (for example, a vacation trip to Yucatan versus Costa Meralda) or several outcome bundles having a probability distribution over the several outcomes (that is, bets). The approach of cognitive psychologists to the study of such decision making is basically descriptive rather than normative. For example, in studying how people evaluate bets, the psychologist will examine how decisions are affected by the importance different individuals assign to amounts of gains or losses, and to the probabilities of wins and losses (see Slovic & Lichtenstein, 1968). Whereas the evaluation of a single "outcome-by-probability" package seems predictable by multiplying the subjective scale values of the outcome and the probability, the total value of two such outcome-by-probability packages is consistently less than the sum of the values of the two packages taken separately (Shanteau, 1974). Of great interest to studies of decision making are recent papers by Kahneman and Tversky (1972, 1973) and Tversky and Kahneman (1971, 1973). They find that when people estimate probabilities of events, they use intuitive heuristics which produce systematic biases in their estimates. One heuristic is "representativeness," wherein people assign higher probability to any outcome that conforms more closely to their stereotype of what a "representative example" ought to look like. Thus, the sequence HTHT will be judged as a more probable outcome of four tosses of a fair coin than will the sequence HHHH, although objectively the two sequences are equally probable. A second heuristic biasing relative frequency judgments is relative availability in memory of the events being judged. Thus, for example, a person will readily generate more words beginning with the letter k than he will words having k in their third position (a fact about memory organization); correspondingly, he will also estimate a higher frequency for the former class of words—which, in fact, is false. The literature cited may be read to see some of the specific biases these heuristics cause when they are applied to various real-world instances of intuitive explanations and predictions.

F. Response Generation

Once an action has been decided on (such as to open the flue of the fireplace), a central plan is drawn from long-term memory for that action and is initialized (or "parameterized") for the current situation in which

the response is to be executed. It appears that most skilled actions may again be analyzed in terms of a multilevel hierarchy with various tests in the plan which compare accomplished with anticipated results. Interestingly, just as we may think of perception as the fitting of a concept onto the current sensory field, so may we think of action in terms of the subject trying to realize a certain concept in his own behavior. Subjects can obviously categorize others' behaviors according to concepts (for example, "he's striking a match," "he's writing his name"). It seems reasonable to suppose that, given an intention, the person then makes those movements necessary under the circumstances to result in his own behavior instancing that concept. Thus, we might think of an action family as those sequences of movements which pass the criteria for being classified under the circumstances as exemplars of a given action concept.

Most contemporary theories of motor organization rely heavily upon concepts from servocontrol theory (see Bernstein, 1967; Pew, 1974). Such theories assume the existence of components necessary for servoregulation of movements: *effectors* that move limbs, *sensors* that provide feedback, a *control element* that contains the current goal stipulated in terms of the desired setting of an intensity or location parameter, a *comparator* that determines the size and sign of the discrepancy between the target value and the achieved value, and a *translator* that converts a given discrepancy into appropriate correctional movements designed to reduce the discrepancy. For instance, in tracking an erratically oscillating spot of light (say, on a TV screen) with one's fingertip, the goal is to maintain a null discrepancy between the momentary location of the spot and the fingertip. To the extent that there are large delays between sensing a discrepancy and firing off a corrective movement the response of the system will lag (be in error) because the dot has moved elsewhere during the delay. The error can be reduced considerably if the movement of the dot is regular and periodic (for example, traces out a sine wave as it sweeps across the TV screen). A higher-level "anticipator" that has learned the periodicity can provide anticipatory commands to the control element moment by moment, so that the fingertip will be moved to where the dot is expected to be in the next instant. It is such predictive knowledge that allows the motor system to move beyond the low-level actions of merely corrective reactions. It is generally supposed that at the higher levels, the solution to a motor problem ("open the flue") is represented in advance in the brain as an abstract motor image of the person's surroundings, the actions to be carried out, the instruments and support relations to be used, and the expected effects of these actions. Such motor programs seem to be using again the basic ideas of hierarchical control: set a goal, select a plan or method, operate for awhile, check your results to see how you are progressing, and so on.

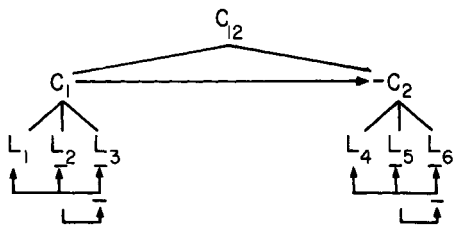

FIGURE 8 Diagram of the associative structure representing a two-chunk series, each composed of three letters. Lines indicate excitatory connections; arrows ending in a minus sign denote inhibitory connections. (Reproduced by permission from Estes, 1972.)

Along somewhat different lines, Johnson (1970) and Estes (1972) have been developing hierarchical models for the organization of serial responses. The typical domain of application is to a human subject learning to smoothly speak a chunked letter series such as HPJ–MXK (see also Lesgold & Bower, 1970). Estes provides an associative theory wherein higher-order "control elements" are introduced to represent chunks, with associations from this element to the items within the chunk. The order of elements is encoded in terms of *inhibitory* links between elements within a chunk. Figure 8 from Estes (1972) shows the internal representation of the series. The *C*'s stand for control elements, the *L*'s stand for the motor programs for uttering the particular letters, lines denote excitatory connections, and arrows ending in "minus signs" denote inhibitory connections. The way the structure fires off is as follows: the decision to respond causes activation of the top control element (C_{12}), so excitation then flows to its next-level elements (C_1 and C_2). But because C_1 inhibits C_2, C_1 will have a greater activation. It passes excitation to its letter tokens L_1, L_2, L_3, of which L_1 occurs first because it is least inhibited. After L_1 occurs, it releases its inhibition of L_2 and L_3. Then L_2 occurs, since it is more activated because it is still inhibiting L_3. After L_2, then L_3 occurs. This completes the chunk C_1, releasing inhibition on chunk C_2, so that it begins popping out L_4, L_5, then L_6. Estes' model for output, is much the same as Johnson's, except for the conjecture regarding *inhibitory* connections (and their release) as encoding serial order.

One advantage of such models, however formulated, is that they can permit very rapid sequential responses which are organized centrally and are not strictly dependent upon moment-by-moment feedback stimulation providing the stimulus for the next response in the chain. In a critique of the early stimulus–response theory of chaining (that is, where feedback from response n was assumed to be the stimulus eliciting response $n + 1$ in the series), Lashley (1951) had objected that this circular mechanism could not operate fast enough to explain very fast sequential responses (for example, a skilled pianist racing through an arpeggio). Chunking models like Estes' solve the problem since all responses within a chunk are activated at once, and each can occur very rapidly as soon as the

inhibitory input preventing its occurrence has been released by the evocation of its predecessor.

Although there is very much more that could be said about motoric processes—their relation to early knowledge representations (Bruner), to early imagery (Piaget), to thinking (Bartlett), and how speech is produced—space limitations prevent our going further into the matter here.

IV. COMMENTS ON LEARNING PARADIGMS

A. Induction

It is fitting in a volume devoted to cognition and learning to end with some brief comments on how the cognitive system sketched in the foregoing applies to the various learning paradigms that have been investigated so intensively by psychologists. Roughly speaking, learning experiments may be viewed as settings which expose the subject to controlled opportunities for acquiring information. This information typically refers to the correlations (arranged by the experimenter) between environmental events or between the subject's actions and environmental outcomes.

The old associationists saw that the human brain was a powerful induction device—a correlation detector—and they explained that ability in terms of its more primitive ability to associate ideas (corresponding to sensations or actions) according to their continguity in time or space. Thus, I associate my easy chair with my fireplace because they occur together in space; I associate striking a match with its lighting because they occur contiguously in time.

Within certain boundary conditions and with emendations and exceptions to be noted, such principles seems to explain the learning that occurs in standard situations. A first simple emendation concerns the historical empiricists' theory of "sensation." They subscribed to the view that the sensory field projected an unstructured mass of sensory elements that caused various features and properties to "light up" in the brain, and hence to become interassociated by virtue of their being experienced contiguously in time and space. That view appears too simplistic, however: rather, the sensory field seems to become organized almost immediately in terms of objects and their interrelations. Gestalt principles of perceptual organization are brought to bear in segregating the field into objects, describing properties as unified with their objects, and identifying objects mainly in relationship to one another. The input to memory can not be considered to be a disorderly mosaic of punctate sensory elements; rather, the perceptual system provides working memory with highly structured contents for storing in LTM.

B. Conditions for Learning

A second emendation is that neither temporal contiguity nor spatial contiguity of *objective* events (or properties) are necessary or sufficient conditions for establishing permanent associations. What is critical is that the two mental representations enter into a single propositional structure in STM; this is the common result of those circumstances which promote "belongingness" of the two events (see Thorndike, 1931). Contiguity without belonging creates only temporary connections which rapidly fade from STM.

Besides the facts about the role of belongingness in learning, other facts suggest that contiguity is not sufficient for learning. One is our failure to learn the millions of series and patterns of sensory events that we experience every day. Incidental learning is notoriously poor (for example, what letters go with what spaces in your telephone dial?). Thousands of bits of structured information enter STM every few hours; most is never used, some is used only temporarily to update a rapidly changing world model (for example, "What down is it in the football game?"), and only a small fraction of the information is ever consolidated into LTM, and that is often of an abstract sort (for example, "our football team won"). The critical circumstances promoting LTM consolidation of contiguous events are not fully known but the question has been intensively investigated. With a human subject, any signal or instruction that informs him that the foregoing information is important and valuable will induce him to rehearse and enrich in some way the symbol structure describing the events and their relations. Instructions in human learning experiments typically serve this role of identifying for the subject exactly what is to be noticed and rehearsed.

With lower animals (and sometimes people) we often must use biologically important stimlui (rewards or punishers) to signal the same message, viz., "this information is valuable, so learn it." Food to a hungry rat is a significant event that introduces a pause or "articulation joint" in an otherwise homogeneous flow of behavioral events. The reward stops the flow momentarily; it essentially signals the central processor to "rehearse the preceding contents of short-term memory." N. E. Miller (1963) called this consolidating rehearsal the "go mechanism," and showed how it could explain a number of findings regarding reward and learning.

The learning mechanism seems to become "switched on" mainly when environmental events do not confirm expectations—when they are surprising or informative. This seems an appropriate rule to follow if a system were to be designed to learn new information but to not waste space storing redundant, predictable information. This assumption regarding informativeness seems supported by the research on the "blocking" of learning

(Egger & Miller, 1962; Rescorla, 1972): in such experiments, if Cue *A* already predicts some important event (say, a painful shock), then trials pairing Cues *A* and *B* before shock will result in no conditioning of Cue *B*. It is as though since *A* predicts shock and that correlation continues during the *A* + *B* trials, there is no new surprising information to be learned, so the *B*-to-shock association is not rehearsed.

The work of Garcia and Koelling (1966) (see Bolles, Chapter 7 of this volume) indicates that the stimuli that are selected out of the record of the past to be associated with an unconditioned reaction depends on what that reaction is. These investigators showed that when a rat becomes sick to his stomach, he associates that selectively with his memory of the most recent novel taste he has experienced, disregarding the myriad auditory and visual stimuli that have passed through STM during the interval between the novel taste and his becoming sick. The result illustrates a selective "mental contiguity" of obvious biological utility (for example, it is most likely that you get stomach sickness from some unfamiliar thing you ate).

Such results suggest that contiguity of events in objective time is not a necessary condition for associating them. Similar selective associations of events across a time span of irrelevant events are easily demonstrated with people. For example, if someone sets my house on fire, I will selectively remember any recent episode in which some enemy threatened to burn my property, and associate (inferentially) the arson with that enemy. As an experimental example, Jacoby (1974) asked some of his subjects to judge whether each successively presented word was in the same semantic category as the just preceding item of the series; other subjects judged whether each item was in the same category as *any* preceding item of the series. In cases where two items of a category (dog–giraffe in the "animals" category) were presented many items apart, the latter subjects showed very much higher recall of one item given the other item as a cue ("dog" given "giraffe"). What was clearly important during the learning trial was the revival of the earlier instance and joint rehearsal of the two items at the time the second item occurred. Such "mental contiguity" was far more important for association formation that was "objective contiguity" of the items. In the Jacoby experiment, instructions designated the terms to be put into mental contiguity; in the Garcia and Koelling experiment, genetic programming dictated the selectivity.

C. Cognitions or Habits?

A perhaps distinguishing feature of the cognitive approach is that it assumes that "what is learned" are conceptual associations encoding knowledge about the relationships among events in the learning situation; this

may be contrasted with the traditional doctrine that a stimulus–response habit is what is learned. Thus, a human adult being trained on an elementary "operant discrimination" problem could detect the relevant correlations and frame for himself "if–then" propositions expressing such contingencies as, "when the high-pitched tone sounds, if I press this lever, I get a dime about every tenth press; if the high-pitched tone isn't on, then I get nothing for pressing." The point of calling that a "conceptual structure" is simply to recognize the versatility of ways by which we may instill, activate, or modify that simple knowledge structure, and the versatility of effective actions the structure will mediate in altered circumstances. For one thing, a human adult could learn by simply being told the contingencies, or by observing another person being exposed to the contingencies. The experimenter may alter details of the exact response (for example, "use your left hand rather than your right"), or payoff in convertible tokens rather than dimes, or convey complete information about altered contingencies (for example, "there'll be no more dimes"), and the subject will quickly alter his actions appropriately. The subject not only presses the lever but can describe the contingencies to himself or to a friend in need of dimes. He becomes "aware of" the contingencies, and only then does he begin to behave appropriately (provided he values dimes and the experimenter's good will). Recent reviews of the critical evidence regarding the role of awareness in human classical and instrumental conditioning (Brewer, 1974; Dulany, 1968) indicate that people put through such procedures figure out the contingencies (either exactly or with a closely correlated hypothesis), and then respond according to their expectations and the outcomes they want to get. "Figuring out the contingencies" simply means becoming aware of or able to label the stimulus–action–outcome correlations programmed into the situation by the experimenter.

Although our examples above use the verbally competent adult, we may suppose that a similarly abstract conceptual proposition encoding the temporal sequence of significant events is learned by nonverbal humans and animals. To quote from Anderson and Bower (1973):

> What complex "abstract structure" is, say, a dog learning when a tone is consistently followed by a painful, noxious stimulus? There is no reason to deny *a priori* that dogs and other nonverbal organisms can form the equivalents of elementary propositions encoding the temporal sequence of significant events. To be absurdly concrete, suppose the dog acquires an associative structure expressible as the proposition: "In experimental situation S, a high-pitched tone is followed within a few seconds by a painful stimulus to my left front paw." If this proposition is combined with other general propositions, such as "If a limb is about to be injured, move it away to avoid injury," the dog could "derive" the command to flex its left paw when the tone sounds. The claim is that such conceptual propositions capture the flavor of the dog's behavior, especially in new situations. (This was Edward Tolman's main position.) Thus, it is known that the

> overt paw-flexion response need not occur during learning. Blockage of the neuromuscular junctures with a curare-like drug prevents all overt responses, yet the tone-shock association may be established under curare, and then later tested with positive results (i.e., conditioned paw withdrawals to the tone) after the drug has worn off. Also the tone–pain association institutes in the dog a totally different "repertoire of responses" to the tone in other situations. Thus, in a later appetitive instrumental conditioning situation, the tone will serve as a "conditioned emotional stimulus" suppressing appetitive behavior. The tone can also be used as a "punishing" event, supporting passive avoidance learning; also, its onset can initiate various kinds of previously learned escape behavior, and its termination can "reinforce" escape behavior. There is, thus, a diverse class of different behaviors affected by the tone-shock experience. Beyond this, there are probably other means to induce a similar memory structure in the dog. For example, the tone could be paired with other painful stimuli like burns, loud noises, nausea, or pinpricks delivered to the left paw.
>
> The tenor of these comments is that for even a nonverbal organism and for a singularly simple connection such as "tone–shock," we still require for its representation an intervening propositional structure which can account for multiple determinants, multiple contexts of retrieval of that information, and multiple varied "behavioral indices" related to that learning [pp. 36–37].

Although such considerations argue that the effects of a conditioning history may be represented in terms of a propositional structure, the information-processing approach itself does not claim to "solve" the many scientific puzzles that surround investigations of conditioning and learning. The theoretical language we use does not automatically tell us when conditioning will occur and when it will fail, how reinforcing events are to be characterized, how parameters of the learning situation exert their influence on the process, or how species-specific action patterns may intrude to override and interact with processes of associative learning. These are genuine puzzles to be investigated regardless of whether one adopts as his theoretical language stimulus–response concepts or information-processing concepts. With that general disclaimer, let us consider briefly the two issues of conditioning rate and forgetting.

D. Contextual Variability and Learning Rate

In analyzing conditioning, we have assumed that the unconditioned stimulus enters STM and also activates rehearsal of the contents of STM. Since these contents become associated to the unconditioned stimulus (US), why is conditioning not always complete in one trial? The answer is that the speed of learning will depend on the variability over trials in the contents of STM and working memory at the time the UCS arrives. Although standard experimental arrangements almost insure that the critical to-be-conditioned element (CS) is in the STM at the time of the US, one cannot

control and keep constant the sensory or ideational contents that are coincidentally filling STM at the start of the trial or during the interval between the CS and the US. This variability in contextual elements in STM acts very much like the trial-by-trial variability of stimulus samples as envisioned in the stimulus sampling theory of Estes (see Estes & Suppes, 1974). The contents of STM on a given trial early in learning will serve as a set of independent retrieval cues; since only a few of these cues will retrieve the prior association to the US, the US will be anticipated with only a low probability. Over trials, the critical elements in the CS–US correlation will be isolated due to variation of irrelevant details.

To illustrate, suppose that on Trial 1 the contents of STM are elements *a*, *b*, *c*, *d* when the UCS occurs; on Trial 2 they are *a*, *b*, *x*, *y*; on Trial 3 they are *z*, *b*, *x*, *k*; and so on. We would like to specify a learning algorithm that will eventually assign a high weight to element *b* in calculating an expectation of the US. A simple algorithm which does so is one that increases the weight (attentional strength or perceptual importance) of an element whenever it occurs with the event to be predicted (US), and which decreases the weight of the element either if (1) the stimulus element occurs without the US, or (2) the element occurs in the presence of a more valid predictor of the US, or (3) the US occurs in the absence of the stimulus element (see Rescorla, 1972). Such an algorithm essentially states a mechanical procedure for the brain to compute trial by trial the current correlation between presence or absence of a cue and presence or absence of the UCS. After learning, then, the most valid cue (the CS) tends to receive high attentional weighting and thus the subject's expectation will be determined by the association retrieved by this most valid cue in the STM during a trial.

We thus view Pavlovian conditioning essentially as "discrimination learning," as the subject learning to isolate the most valid predictive cue (of the US) in a variable environment. Operant conditioning would be viewed similarly except the important element entering STM would be feedback from (or the efferent command issued for) the response upon which reward is contingent. That is, the subject would learn a correlation of the form, "in situation *S* response *R* is followed by outcome *X*." Such knowledge modules are then retrieved and activated whenever the subject is in a situation resembling *S* and has a present demand or desire for outcome *X*.

E. Retention of Learned Associations

The primary factors contributing to retention losses of well established associations are contextual changes in the prevailing stimuli and possibly the learning of competing events or responses to the stimulus during the retention interval. The chances of these negative factors occurring increases over

the usual retention interval. Contextual changes refer to alterations over time in the setting and thoughts that fill STM or the direction of attention (to different sensory dimensions), so that an alered pattern of stimulation serves as the retrieval compound. The compound is less likely to retrieve the designated association the greater the contextual change (see Bower, 1972, or Estes, 1955, regarding the "fluctuation" model of contextual alterations and their effect on retention).

The second factor promoting forgetting, namely interfering learning, has been much studied, and the basic principles seem to apply quite generally to various paradigms, responses, and species (see Postman, 1971). If the subject learns an association *A–B*, then to the cue of *A* he is likely to forget it or to become confused in remembering if he has also learned the competing association *A–C* either *before* or *after* the *A–B* learning. Cognitive theory has no distinctive hypothesis to offer regarding the underlying causes of proactive or retroactive interference; they are simply accepted as phenomena in need of explanation. Anderson and Bower (1973, Chapter 15) offered the interpretation that the events *A–B* were acquired by elaboration of a proposition (say *A* as subject, *B* in the predicate), that *A–C* was acquired similarly, and that this results in a temporally ordered stack of associations out of *A*. An inherent restriction on retrieval times would then produce the observable phenomena of retroactive interference, whereas restricted retrieval plus a random reordering of the association stack out of *A* would produce the observable phenomena of proactive interference. The references cited should be checked for arguments for and against this hypothesis.

F. Cognitive Psychology and Social Learning Theory

While commenting briefly on the way cognitive psychology applies to some laboratory learning task, it is appropriate to note that cognitive psychology helps us also understand the learning and utilization of so-called "personality" and social behaviors. Its applications to socialization and personality development are at least as cogent, if not more so, than the account provided by S–R reinforcement theory. To cite but one example, there has been a continuing complaint that the so-called "behavior therapies" or "behavior modification procedures" use techniques of diagnosis, persuasion, motivation, imagery, observational modeling, thought control and image control, etc. in their therapy, despite the fact that none of these techniques are at all explicable in terms of the basic framework of behaviorism and conditioning (see Breger & McGaugh, 1965; Locke, 1971). The ascendant if not prevailing viewpoint today in accounts of socialization, self-control, and personality development is "social learning theory" (e.g., Bandura 1969, 1971; Mischel 1969, 1973). It differs from S–R reinforcement theory in its emphasis on mechanisms of observational learning,

self-control, and "cognitive factors" in the acquisition and utilization of social behaviors. Social learning theorists refer repeatedly to "internal representations" of observed situation–action–outcome sequences (for example, in language codes, in images, or in conceptual propositions), and how these are remembered, utilized, and guided by anticipation of the desired outcome. It seems that cognitive psychologists are studying those intellectual skills that social learning theorists feel compelled to refer to repeatedly in explaining how a person acquires knowledge of his social environment and acquires skill in negotiating his way through it. Thus, amalgamation of cognitive psychology with social learning theory should provide a scientific but broadly relevant synthesis of a sort that psychology has so long been searching for.

ACKNOWLEDGMENTS

The writing of this chapter was supported by a research grant, MH-13950, from the National Institutes of Mental Health.

REFERENCES

Abelson, R. P. The structure of belief systems. In R. C. Schank & K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: Freeman, 1973.

Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D.C.; Winston, 1973.

Anderson, R., & Ortony, A. On putting apples into bottles: A problem of polysemy. *Cognitive Psychology*, in press.

Appel, L. F., Cooper, R. G., McCarrell, N., Sims-Knight, J., Yussen, S. R., & Flavell, J. H. The development of the distinction between perceiving and memorizing. *Child Development*, 1972, **43**, 1365–1381.

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*, August, 1971, 82–90.

Bandura, A. *Principles of behavior modification*. New York: Holt, Rinehart & Winston, 1969.

Bandura, A. *Social learning theory*. New York: General Learning Press, 1971.

Bartlett, F. C. *Remembering*. Cambridge, England: Cambridge University Press, 1932.

Belmont, J. M., & Butterfield, E. C. Learning strategies as determinants of memory deficiences. *Cognitive Psychology*, 1971, **2**, 411–420.

Bernstein, N. *The Co-ordination and regulation of movements*. Oxford: Pergamon Press, 1967.

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.

Blakemore, C. Development of the mammalian visual system. *British Medical Bulletin*, 1974, **30**(No. 2), 152–157.

Bower, G. H. A descriptive theory of human memory. Paper given at the Second Annual Princeton Conference on Learning and Remembering, Princeton, New Jersey, 1964. Proceedings published in D. Kimble, (Ed.), *Learning, memory and forgetting*. Vol. 2, New York: New York Academy of Science, 1967.

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.

Bower, G. H., & Winzenz, D. Comparison of associative learning strategies. *Psychonomic Science*, 1970, **20**, 119–120.

Breger, L., & McGaugh, J. L. Critique and reformulation of "learning theory" approaches to psychotherapy and neurosis. *Psychological Bulletin*, 1965, **63**, 338–358.

Brewer, W. F. There is no convincing evidence for operant and classical conditioning in humans. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and Symbolic Processes*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1974.

Broadbest, D. E. *Perception and communication*. New York: Pergamon, 1958.

Charniak, E. Toward a model of children's story comprehension. Technical Report AI TR-266. Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1972.

Conrad, R. Acoustic confusions in immediate memory. *British Journal of Psychology*, 1964, **55**, 75–84.

Craik, F. I. M. A "levels of analysis" view of memory. In P. Pliner, L. Krames, & T. Alloway (Eds.), *Communication and affect: Language and thought*. New York: Academic Press, 1973.

Crothers, E. J. Memory structure and the recall of discourse. In J. B. Carroll and R. O. Freedle, (Eds.) *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston, 1972.

Duda, R. O., & Hart, P. E. *Pattern classification and scene analysis*. New York: Wiley–Interscience, 1973.

Dulany, D. E., Jr. Awareness, rules and propositional control: a confrontation with S–R behavior theory. In D. Horton & T. Dixon (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

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.

Ellis, N. R., Detterman, D. K., Runcie, D., McCarver, R. B., & Craig, E. N. Amnesic effects in short term memory. *Journal of Experimental Psychology* 1971, **89**, 357–361.

Epstein, W. Mechanisms of directed forgetting. In G. H. Bower (Ed.), *The psychology of learning and motivation*. Vol. 6. New York: Academic Press, 1972. Pp. 147–192.

Estes, W. K. Statistical theory of spontaneous recovery and regression. *Psychological Review*, 1955, **62**, 145–154.

Estes, W. K. Reinforcement in human learning. In J. Tapp (Ed.), *Reinforcement and behavior*. New York: Academic Press, 1969.

Estes, W. K. An associative basis for coding and organization. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972. Pp. 161–190.

Estes, W. K. Phonetic coding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 360–372.

Estes, W. K., & Suppes, P. Foundations of stimulus sampling theory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & (Eds.), *Contemporary developments in mathemati-*

cal psychology. Vol. I: *Learning, memory, and thinking*. San Francisco, California: Freeman, 1974.

Feigenbaum, E. A. Stimulation of verbal learning behavior. In E. A. Feigenbaum & J. Feldman, (Eds.), *Computers and thought*. New York: McGraw-Hill, 1963.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. Developmental changes in memorization processes. *Cognitive Psychology*, 1970, **1**, 324–340.

Frederiksen, C. H. Effects of task-induced cognitive operations on comprehension and memory processes. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and acquisition*. Washington, D.C.: Winston, 1972.

Garcia, J., & Koelling, R. A. Relation of cue to consequence in avoidance learning. *Psychonomic Science*, 1966, **4**, 123–124.

Garner, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.

Gibson, E. J. *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts, 1969.

Glanzer, N., & Clark, W. H. Accuracy of perceptual recall: An analysis of organization. *Journal of Verbal Learning and Verbal Behavior*, 1963, **1**, 289–299.

Glass, A. L., & Holyoak, K. J. Alternative conceptions of semantic memory. *Cognition*, in press.

Glass, A. L., Holyoak, K. J., & O'Dell, C. Production frequency and the verification of quantified statements. *Journal of Verbal Learning and Verbal* Behavior, 1974, **13**, 237–254.

Guzmán, A. Decomposition of a visual scene into three-dimensional bodies. In A. Grasselli (Ed.), *Automatic interpretation and classification of images*. New York: Academic Press, 1969.

Hubel, D. H., & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, 1962, **160**, 106–154.

Hubel, D. H., & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology (London)*, 1968, **195**, 215–243.

Jacoby, L. L. Encoding processes, rehearsal and recall requirements. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 302–310.

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.

Johnson, N. F. The role of chunking and organization in the process of recall. In G. H. Bower (Ed.), *The psychology of learning and motivation*. Vol. 4. New York: Academic Press, 1970.

Kahneman, D., & Tversky, A. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 1972, **3**, 430–454.

Kahneman, D., & Tversky, A. On the psychology of prediction. *Psychological Review*, 1973, **80**, 237–251.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1974.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley, 1951.

Lesgold, A. N. & Bower, G. H. Inefficiency of serial knowledge for associative responding. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 456–466.

Lindsay, P. H., & Norman D. A. *Human information processing: An introduction to psychology*. New York: Academic Press, 1972.

Locke, E. A. Is "behavior therapy" behavioristic? *Psychological Bulletin*, 1971, **76**, 318–327.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation.* Vol. 1. New York: Academic Press, 1967.

Maslow, A. H. A theory of metamotivation: The biological rooting of the value life. *Journal of Humanistic Psychology,* 1967, **7,** 93–127.

Meacham, J. A. The development of memory abilities in the individual and in society. *Human Development,* 1972, **15,** 205–228.

Meyer, B. J. F. The organization of prose and its effect on recall. Technical Report No. 1 (Ph.D. dissertation) (Reading and Learning Series), Department of Education, Cornell University: Ithaca, New York, 1974.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 1956, **63,** 81–97.

Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, Rinehart & Winston, 1960.

Miller, N. E. Some reflections on the law of effect produce a new alternative to drive reduction. In M. R. Jones (Ed.), *Nebraska symposium on motivation.* Lincoln: Nebraska University Press, 1963. Pp. 65–112.

Mischel, W. Continuity and change in personality. *American Psychologist,* 1969, **24,** 1012–1018.

Mischel, W. Towards a cognitive social learning reconceptualization of personality. *Psychological Review,* 1973, **80,** 252–283.

Morton, J. The interaction of information in word recognition. *Psychological Review,* 1969, **76,** 165–178.

Newell, A., *et al. Speech understanding systems: Final report of the study group.* New York: American Elsevier, 1973.

Newell, A., & Simon, H. A. *Human problem solving.* Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Norman, D. A., & Rumelhart, D. E. *Explorations in cognition.* San Francisco: Freeman, 1975.

Parkinson, S. R., Parks, T. E. & Kroll, N. E. A. Visual and auditory short-term memory. *Journal of Experimental Psychology,* 1971, **87,** 274–280.

Pew, R. W. Human perceptual-motor performance. Technical Report No. 48. Human Performance Center, University of Michigan, Ann Arbor, Michigan, August, 1974.

Postman, L. Transfer, interference and forgetting. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's experimental psychology.* (3rd ed.) New York: Holt, Rinehart & Winston, 1971.

Prytulak, L. S. Natural language mediation. *Cognitive Psychology,* 1971, **2,** 1–56.

Ratliff, F. *Mach bands: Quantitative studies on neural networks in the retina.* San Francisco: Holden-Day, 1965.

Reed, S. K. Pattern recognition and categorization. *Cognitive Psychology,* 1972, **3,** 382–407.

Rescorla, R. A. Informational variables in Pavlovian conditioning. In G. H. Bower (Ed.), *The psychology of learning and motivation.* Vol. 6. New York: Academic Press, 1972. Pp. 1–46.

Rieger, C. J. Conceptual memory: A theory and computer program for processing the meaningful content of natural language utterances. Unpublished doctoral dissertation, Stanford University, 1974.

Rumelhart, D. E. Notes on a schema for stories. Unpublished working paper, Psychology Department, University of California at San Diego, October, 1973.

Rundus, D. Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology,* 1971, **89,** 63–77.

Schank, R. C. Conceptual dependency: a theory of natural language understanding. *Cognitive Psychology*, 1972, **3**, 552–631.

Schank, R. C. The development of conceptual structures in children. Technical Report AIM-203, Artificial Intelligence Laboratories, Computer Science Department, Stanford University, May 1973. (a)

Schank, R. C. Identification of conceptualizations in natural language. In R. C. Schank & K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: Freeman, 1973. (b)

Shanteau, J. Component processes in risky decision making. *Journal of Experimental Psychology*, 1974, **103**, 680–691.

Shiffrin, R. N., & Atkinson, R. C. Storage and retrieval processes in long-term memory. *Pshchological Review*, 1969, **76**, 179–193.

Simon, H. A. Motivational and emotional controls of cognition. *Psychological Review*, 1967, **74**, 29–39.

Skinner, B. F. *About behaviorism*. New York: Knopf, 1974.

Slovic, P., & Lichtenstein, S. Relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Psychology Monograph*, **78**, (No. 3, Part 2), 1968.

Sternberg, S. The discovery of processing stages: extensions of Donder's method. Attention and Performance. *Acta Psychologica*, 1969, **30**, 276–315.

Taylor, C. *The Explanation of Behavior*. London: Routledge & Kegan, Paul, 1964.

Thorndike, E. L. *Human learning*. New York: Century Co., 1931. (Paperback ed., Cambridge, Massachusetts: MIT Press, 1966)

Tolman, E. C. *Purposive behavior in animals and men*. Berkeley: University of Calfornia Press, 1932.

Tulving, E. Subjective organization and effects of repetition in multi-trial free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 193–197.

Turvey, N. T. Peripherial and central processes in vision: Inferences from an information processing analysis of masking with pattern stimuli. *Psychological Review*, 1973, **80**, 1–52.

Tversky, A., & Kahneman, D. Belief in the law of small numbers. *Psychological Bulletin*, 1971, **76**, 105–110.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, **5**, 207–232.

Tversky, B. Encoding processes in recognition and recall. *Cognitive Psychology*, 1973, **5**, 275–287.

Vicens, P. Aspects of speech recognition by computer. Technical Report CS-127, doctoral thesis, Computer Science Department, Stanford Univ., 1969.

von Békésy, G. *Sensory inhibition*. Princeton, New Jersey: Princeton University Press, 1967.

von Restorff, H. Analyse von Vorgangen in Spurenfeld. I. Über die Wirkung con Bereichsbildungen im Spruenfeld. *Psychologisch Forschung*, 1933, **18**, 299–342.

Waugh, N. C., & Norman D. A. Primary memory. *Psychology Review*, 1965, **72**, 89–104.

Wescourt, K. T., & Atkinson, R. C. Scanning for information in long- and short-term memory. *Journal of Experimental Psychology*, 1974, **98**, 95–101.

Winston, P. H. Learning structural descriptions from examples. MIT Artificial Intelligence Laboratory. Project AI-TR-231, 1970.

3

Cognitive Theory Applied to Individual Differences

Earl Hunt
Marcy Lansman

The University of Washington

All men may be created equal, but no two are created identical. Over a hundred years ago Sir Francis Galton accepted this fact as a central problem for psychology. He devoted much of his life to an attempt to explain individual differences in mental ability in terms of general psychological laws. For a number of historical and social reasons Galton's goal of scientific explanation was put aside in order to develop a sophisticated system for measuring cognitive power without attempting to explain it (Tyler, in press). The result is the intelligence test that we know today, and which has become our normal way of describing differences in cognition. Yet the questions Galton posed are important scientific issues. We shall try to restate them in terms of the modern theory of cognitive psychology.

What might constitute a theory of individual differences? Since the early 1900s, great emphasis has been placed upon classification of the basic ways in which individuals' thought processes differ. Classification is certainly a legitimate stage in the development of a scientific theory. However, we part company with the traditional theory of mental testing, or *psychometrics*, in stating that we do not regard a classification scheme as an adequate final goal. Instead we want a theory which describes the *processes* by which different people attack the same problem.

To illustrate what is meant, consider an analogy to the measurement of automobile abilities. An observer could develop a classificatory scheme

by noting that vehicles differed in their weight, speed, gasoline consumption, turning ratio, and so forth. By appropriate statistical analysis these data could be reduced to a small number of inferred dimensions. For instance, it would soon be noticed that weight, acceleration, and gasoline consumption covaried, and such an observation might lead to the derivation of an abstract concept of *power*. Certainly such a concept is useful in thinking about automobiles, but for a really adequate description of how a car works we use our knowledge of engine design, and of the workings of steering and transmission systems. A listing of vehicle characteristics just is not an adequate way to explain how an automobile negotiates a turn—or fails to.

We find equally unsatisfactory the statement that an individual is, or is not, suited for a particular position "because of his or her IQ." In order to understand differences in intellectual performance we must know how that performance occurs, and how specific individuals deviate from the norm.

This sort of reasoning would certainly have pleased Galton, but it is not the dominant philosophy in psychometrics today. We have already noted that psychometricians basically rely on a classificatory approach. The mechanics of their approach, however, are quite sophisticated. Let us give a brief, nonmathematical introduction. For further details, see Edwards (1970) or, for a complete technical discussion, Harman (1967).

Suppose that a battery of mental tests are administered to a diverse group of people. The tests themselves are generally different, but all are agreed to require mental processes. Typical examples are tests of vocabulary, rote memory, mathematical computation, and the visualization of spatial relations. We would find that scores on most such tests are positively correlated; that is, those who do well on one tend to do well on others, but the rankings are far from identical. Further, we would find that certain tests seem to cluster together, so that the correlations between tests within a cluster are higher than the correlations between pairs of tests in different clusters. How are we to interpret such results?

Factor analysis is a mathematical technique that has been developed to analyze sets of correlations. Actually, it is not one technique, but a set of computational methods, each of which makes slightly different assumptions, and each of which may yield a different answer. All factor-analytic techniques share the assumption that an individual's performance on each test is due to the weighted sum of a relatively small number of basic underlying abilities. The weighting will differ for each task, but the principle is always the same—performance is due to a summation of abilities. The factor-analytic techniques are designed to determine how many dimensions (underlying abilities) and what weights are needed to account for the observed pattern of correlations between tests.

Factor analysts originally assumed that a rather small number of underlying abilities would be adequate to explain a large part of the variance in virtually any battery of cognitive tests. While this hypothesis has not been rejected, the picture now appears to be much less clear. Spearman (1927) argued that the data could be explained by assuming a single general ability (g) with the addition of a specific factor for each type of test. This position has since been developed further by Vernon (1961) and Cattell (1971). Roughly, their argument is that there exists a hierarchy of abilities—a single general factor, plus subsidiary factors of verbal, numerical, and spatial ability, and subfactors of each of these. In contrast to this position is the position that the mental abilities, instead of being arranged in a hierarchy, represent different, but equally general, dimensions of a Euclidean space of cognitive performance. Thurstone (1938), the leading exponent of this view, believed that seven distinct abilities could be identified. More recently, Guilford (1967; Guilford & Hoepfner, 1971) has asserted that at least 120 separate abilities can be identified.

Many of the disparities between these two camps result from disagreements on the technique of factor analysis. It does not seem too harsh to say that each factor analyst makes an assumption about the basic nature of the mental space (for example, it is Euclidean or hierarchical) and then uses a data-reduction technique that identifies the best underlying variables *within* the class of solutions permitted by the original assumption.

We do not mean to imply that all is confusion. Most psychologists would agree that describing a person by a *single* intelligence score is not sufficient. (Even so, it is amazing how many studies are still published which refer to mental power in terms of the simplistic IQ score.) In contrast, it is seldom necessary to study Guilford's 120 abilities. There is strong evidence for three distinct mental abilities: language comprehension, reasoning, and the visualization of spatial relations. Verbal tasks are heavily influenced by whatever the language comprehension score measures, and mathematical reasoning appears to be a combination of general reasoning and spatial visualization ability. Since verbal and mathematical ability are important in many academic and professional situations, most widely used academic screening tests provide "verbal ability" and "mathematical ability" scores. Such scores are impressive predictors of academic success on a population wide basis, although there is ample room for exceptions to the general rule that high test scores are predictive of high grades (Tyler, 1974).

We conclude that psychometric tests provide a reasonable basis for classifying people by mental ability. They are of little use in providing the desired process description of individual cognition. Certainly a person does not do poorly in mathematics because he has obtained a low score on a mathematical ability test. Rather, it must be that test and problem-solving

performance are measures of related thought processes. The correlation between the test and later performance cannot tell us what these processes are.

Why, then, has the psychometric approach to individual differences been predominant? Tyler (in press) has pointed out that classification measures are quite adequate predictive tools, as long as one does not care why the prediction works. For example, the correlation between high school grades and the verbal score of the Wechsler Adult Intelligence Scale (WAIS) is .63 (Matarazzo, 1972). Linguistic comprehension scores are the best single predictors of success in college (Lanvin, 1965). Given the unreliability of grade assignments, these statistical facts are *in themselves* sufficient evidence to warrant the use of the tests to predict academic success. The original motivation for developing intelligence tests was to make such predictions. Binet was asked to construct a test that would locate those individuals who probably could *not* benefit from standard public school instruction, in order that the state be relieved of the obligation to educate children of low mentality up to the point of normal performance. Similarly, the first large group intelligence test, the *Army Alpha* of World War I, was intended solely to identify those citizens who were not likely to be able to absorb standard military training. In both cases the psychometrician was asked "Who can?" and not "How do they?" Binet and his followers answered the question they were asked.

The traditional psychometric measures have also provided basis for studying the nonpsychological correlates of intellectual development. Very many studies have examined the correlations between intelligence test scores and social or biological variables presumed to affect mental ability. Most of this research has been directed at the "nature–nurture" question: What are the relative contributions of genetic and environmental factors to variation in intelligence? Note that we are once again faced with correlational findings; emphasis is on the relative size of biological or social correlates of intelligence, and not on the mechanism by which these variables may influence mental abililty. We could not begin to review the voluminous literature. Appropriate surveys have been written by Matarazzo (1972) and Tyler (1974). We shall comment on the logic of such studies.

The argument for determining the correlation between an intelligence test score and some antecedent variable is that the test score represents a fundamental measure of human ability, which may be produced by the social or biological antecedent. Similarly, in other studies it may be more sensible to regard intelligence itself as the antecedent. Introductory students are correctly taught "correlation does not prove causation." In order to show that an antecedent produces a consequent we must specify some mechanism of action. The fact that there are statistical associations between

intelligence and genetic constitution, and between intelligence and scholastic achievement do not tell us *how* genetic factors determine cognitive capability, or how cognitive capability determines school performance. Furthermore, in correlational studies there is always the problem of the hidden variable. The best example of this is the problem of studying genetic constitution and intelligence. In modern society it is impossible to separate inheritance, socioeconomic status, and education. All of these variables correlate with intelligence test scores and with each other, making the question of causation obscure, to say the least.

Because of such problems, there have been suggestions that intelligence, as defined by traditional psychometric tests, is of no scientific interest. It is certainly true that no one study of the relationship between test scores and any other variable is unchallengeable. Taken together, there are so many studies showing relationships between psychometric intelligence and other important variables that we think the literature cannot be ignored. Modern intelligence tests probably do identify individuals who exhibit marked variation in cognitive power. We should try to find out what the nature of this variation is.

Most of the remainder of our presentation is an elaboration of this theme. One more comment on the methodology of correlational studies is in order. Very many of these studies use a single global measure of intelligence. Indeed, to avoid circumlocution, we have referred to a single variable called "intelligence." This is wrong. Psychologists *know* that intelligence is not a unidimensional variable. Different components of intelligence may be differentially sensitive to particular environmental and social variables, or even follow different biological models. The verbal comprehension and spatial visualization components of intelligence tests provide an excellent example. Bock (1973) has observed that the statistics on the distribution of *verbal* reasoning ability across related individuals suggest that the genetic component of this trait is due to the action of a large number of genes. In contrast, the genetic component of *spatial* reasoning may possibly be due to the action of a single gene located on the X chromosome, and hence may be more likely to be expressed in males than females (Bock & Kolakowski, 1973; Yen, 1973). Similarly, Lesser, Fifer, and Clark (1965) found that different ethnic groups in New York City had characteristic patterns of relative high or low scores on various types of intelligence tests. Orientals typically scored higher on numerical and spatial reasoning tests than on verbal tests, whereas Jews showed the opposite pattern. Because of variation in social factors, this observation cannot be taken as evidence for genetic differences, but it can be taken as a demonstration that the different types of cognitive performance required on the various tests used by Lesser and his co-workers were differentially affected by the social and/or genetic differences between the ethnic groups. We

hope that in the future there will be fewer studies of "the effect of x on intelligence" or "the effect of intelligence on y," and more studies which specify the particular type of intelligence in question.

I. THE INFORMATION-PROCESSING APPROACH

A. Some General Concepts

We are clearly calling for a new approach. What is it? We first introduce some general concepts from the information processing view of psychology, and indicate how they might be used in the study of individual differences. Next we illustrate the approach by considering a series of studies of a situation involving memorization of material for a relatively brief period of time. We then describe a more general model of human cognition, which includes perception, memory, and problem solving. Finally, we describe data relevant to the model in each of these areas.

When we say that we regard man as an information-processing system, in the same sense that a computer system is an information-processing system, we are *not* advocating a digital computer as a physical model of the human brain. What we want to do is to explain human thought using some of the concepts used to describe and analyze computer systems. One such concept is the distinction between *hardware, system architecture,* and *programs* or *control processes.* "Hardware" refers to the physical components used to hold, transform, and transmit information. "System architecture" refers to the way in which different functional components are attached, thus determining how information must flow. The architecture is considered fixed, but usable in different ways, in much the same way that a highway system is fixed, but permits some choice of routes. To take a more physchological example, Sperling (1960) and others have argued that visual information is impressed upon a brief *iconic storage system,* which holds information about the physical characteristics of the stimulus for fractions of a second. Part of this information is read into more central locations for further processing. The iconic store itself is a system architecture concept. The decision to read information from iconic store in a particular order, however, represents a programming or control-process concept. That is, it is one of several alternative rules which can be used to decide what information is selected for entry into certain routes in the information processing system, with the range of selection being established by the system architecture.

Although the physical principles of information processing are undoubtedly the same for all members of the species, different individuals

may have different quality components. We assume that the system architecture is universal throughout mankind. Component quality variation and system architecture constancy can combine to produce important individual differences in the parameters required to describe functioning at the system level. An information-processing system contains three basic types of devices: information-transforming units, memory units, and transmission pathways. A small memory unit or a limited pathway located at a key point could alter system capacity in much the same way that a small bridge can alter traffic flow in a city. Thus, it seems reasonable to expect some fairly permanent individual differences to appear in the parameters which describe components of the system.

Large individual differences can be expected in control processes, both across subjects and within subjects under different experimental conditions. There is a temptation to regard all control processes as learned, and thus relatively flexible. This conception may be inaccurate, since it is possible that there are general human tendencies toward the use of certain types of control processes. Cross-cultural research, which can be considered an extreme variation of an individual differences study, offers some interesting examples. Beginning with studies by Collins and Quillian (1969) a number of experiments have provided evidence that human semantic memory is organized into hierarchical categories. Thus, "canary" is filed in our brains as a species of "bird." (For a review of the evidence, see Millward, Rice, & Corbett, 1975.) In the strictest computer science sense, *both* the use of a hierarchical structure and the particular hierarchies used would be considered control processes. Cole, Gay, Glick, and Sharp (1971) found evidence for hierarchical structuring of information by the Kpelle people of Liberia, but, not surprisingly, the Kpelle use different categories than do undergraduates in the United States. The caution should be clear. Undoubtedly many of the individual differences in our thinking are due to different learning experiences. We should not be fooled into believing that because a particular information handling process would be constructed by program selection in a computer, it therefore follows that the analogous process is a learned one in man.

B. An Example

To illustrate how these concepts may be used, we will describe several studies based upon Atkinson and Shiffrin's (1968) proposal for a model of memory. The architecture of their model is simple. It assumes a short-term memory (STM) that holds at most r items, with r on the order of two to seven. Items are maintained in STM by a process of active rehearsal, so that a person can drop an item from STM either by replacing it with a new item or by failing to rehearse it. The model also contains a long-term

memory (LTM) component that is of conceptually infinite capacity. Although LTM cannot be filled beyond capacity, information can be lost from LTM by some unexplained process. The probability of an item's being retrieved from LTM on demand is assumed to decay exponentially to zero, as a function of the presentation–retrieval interval.

Atkinson and Shiffrin propose that the following events occur when a piece of information is presented to a subject:

1. The information is entered into STM with probability α, which might vary for particular types of stimuli, and is under the subject's control.
2. When an item is entered into STM, another item may drop out. Whether this happens or not depends upon the strategy the subject uses for item rehearsal and upon the number of items already in STM. The latter, in turn, is a function of the size, r, of the STM.
3. Information about an item in STM is transferred to LTM at rate θ. Thus, the total amount of information about an item which is placed in LTM will be a function of θ and of the time the item spends in STM.
4. Information in LTM "decays" (that is, becomes unavailable) by a negative exponential function, with decay parameter τ.

The strategy the subject chooses thus determines how the architecture is to be used. The strategy and the experimental conditions imply a set of equations relating stimulus conditions to response functions. The values of α, r, θ, and τ enter into these equations as parameters, and may themselves be affected by a choice of strategy.

Atkinson and Shiffrin evaluated their model by comparing its predictions to behavior on the *continuous paired-associates task.* In this task the subject is presented with a sequence consisting of arbitrary item associations and queries concerning the latest value of each association. A possible sequence is

$$A\text{–}12, \qquad B\text{–}42, \qquad C\text{–}16,$$

thus establishing initial associates for A, B, and C. After this has been done each trial consists of a query, in which the subject must indicate the response term for an associate, and an update, in which a new response is associated with the stimulus term. Continuing the example, the next trial might be B–? for which the correct response would be "42," followed by a new association B–73. Performance is indicated by plotting the probability of a correct response as a function of *lag,* the number of trials between the last presentation of an update of the stimulus term and the current query. In the example the lag is one.

Atkinson and Shiffrin assumed both a system architecture and a particular strategy for the continuous paired-associates task. From these they derived equations that could be used, in combination with gross features of

TABLE 1
Parameter Estimates for Students of
High and Low Quantitative Ability

Parameter	α	r	θ	τ
High quantitative ability students	.44	1.70	.30	.74
Low quantitative ability students	.27	1.42	.30	.41

the data, to find those values of the parameters that most closely matched the data. They then attempted to predict the fine-grain structure of the data, including variations in probability correct as the experimental conditions were changed. In general they were quite successful. We shall now consider some situations in which the Atkinson and Shiffrin model has been used to study individual differences in the capacity to retain information for several minutes.

Hunt, Frost, and Lunneborg (1973) contrasted the parameters obtained from college students who had either (relatively) high or low scores on tests of verbal comprehension or quantitative ability. They found that individual parameter estimates did not vary as a function of verbal ability, but they did find differences associated with quantitative ability. Table 1 shows the parameter estimates obtained from students with high or low quantitative ability. The differences between the estimates of α and τ are substantial.[1] The minimal statement that can be made is that students with varying psychometrically defined quantitative ability differ in their use of their memory systems. The inference that the "high quantitative ability" subjects are in some way "component superior" is not warranted. The most precise statement is that the data indicates that students with high quantitative ability are (*a*) more likely to enter a new item into STM and (*b*) less likely to be able to retrieve an old item from LTM. This pattern could be produced by a strategy which banks on retrieving items from STM rather than LTM, or, of course, it could be a component difference. While we still do not know which explanation is true, the parameter estimation approach alone has taken us beyond the simple statement that some people have better memories than others, and has given us a qualitative insight into differences in the way which certain subjects use their memory systems.

[1] Hunt, Frost, and Lunneborg (1973) report a significance level of .10 for the differences in α and .002 for τ. Carroll (personal communication) pointed out some inconsistencies in the published data. After reexamining the original scores, Hunt and Lunneborg determined that the correct significance levels were .01 and .001, respectively.

There can be advantages to using a task for which there is an explicit model in the study of individual differences, even if the analysis is not carried to the point of parameter estimation. Love, working in our laboratory, has conducted a number of experiments using *mnemonists* (*M*s, in Love's terms), people who have demonstrated unusual memory abilities at a more or less anecdotal level. The most spectacular of the *M*s, VP, is a man who can astound people with his memory in demonstrations. For instance, he can play several games of chess blindfolded or recall the details of a story read over a year earlier (Hunt & Love, 1972). Love (1973) located several other *M*s who, though not as impressive as VP, were intermediate between him and university students in their ability to do memory tasks. Love used the continuous paired-associates task to gain some insight into how the *M*s achieved their superiority. The results, shown in Figure 1, indicated that the mnemonists were much superior in their ability to recall rapidly presented information for a period of several minutes. In fact, VP was perfect! Love wondered whether the mnemonists' ability to recall information over long periods of time might be due to superior immediate memory and coding, rather than superior retention or retrieval processes. To test this he utilized a paradigm developed by Nelson (1971), in which subjects first learned a list of arbitrary number–noun pairs to a criterion of one perfect recitation. They then relearned the list four weeks

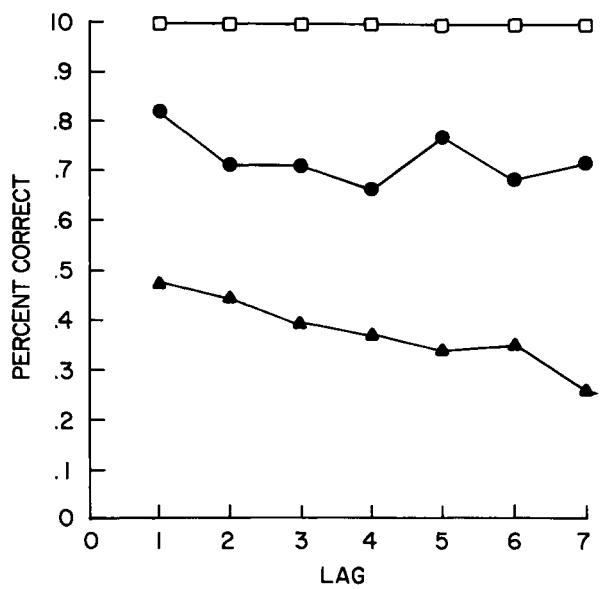

FIGURE 1 Percent correct recall in the Atkinson and Shiffrin (1968) task, as a function of lag: (□) for VP; (●) the 5 mnemonists (*M*s); (▲) 39 unselected college subjects (normals).

later. The *M*s exhibited no better recall of the list than did Nelson's college student subjects. From these and other data, Love concluded that very superior memory is primarily associated with information retention and coding at the time the information is stored, and is not due to a superior capacity to retain information in a quiescent form over a long period of time.

The Atkinson–Shiffrin model has similarly been applied at the other extreme of memory ability, to study qualitative differences in memory functioning of mental retardates (Brown, Campione, Bray, & Wilcox, 1973; Brown, Campione, & Murphy, 1974). This application illustrates some of the difficulties as well as the potential of the information processing approach. A straightforward way to apply the model would be to run a continuous paired-associates experiment using educable mentally retarded subjects (IQ 50–85) and then estimate the parameters. This tactic is not possible, because the task as presented to college students is beyond the attention capacity of the mentally retarded. Brown and her co-workers used a continuous paired-associates task with stimuli modified for retardates. Parameter estimates are no longer comparable to those obtained using the normal version of the task.[2] Therefore, Brown and her co-workers relied upon experimental manipulations to uncover processing strategies. They found that training subjects to use an explicit rehearsal strategy facilitated the performance of retardates, but did not modify the performance of normal subjects (Brown *et al.*, 1974). Conversely, introducing procedures that interfered with rehearsal caused a deterioration in the performance of normals, but not of retardates. They concluded that the difference between retardates and normals in immediate-memory tasks was not in the structure of the memory system, but in the tendency to adopt the active rehearsal strategy. It is of interest to note that Belmont and Butterfield (1969) reached the same conclusion on the basis of a literature review of studies of the relationship between STM and intelligence. While it is accurate to say that retardates have "poor memory," their memory is poor in a particular way. The difference between retardates and normals in memory for ongoing events appears to be due to the use of different strategies for encoding information.[3]

[2] This represents a major problem. The information processing approach dictates the use of tasks which, on occasion, may be administratively difficult to use outside of the typical psychological laboratory.

[3] Retardation is a gross diagnostic category. The subjects used by Brown and her co-workers, and by many others in this field, probably did not represent a homogeneous population with respect to the etiology of their retardation. It would be of interest to know if specific types of memory disoreder are associated with specific etiological categories of retardation, such as Down's syndrome ("mongolism") or any of a number of other specific biological defects which produce retardation as one of several symptions.

C. A More General Model

Cognition involves more than just memory. No general model has been developed for all cognitive phenomena, and we certainly cannot produce one here. We shall discuss a general approach toward the classification of cognitive abilities. This approach is a combination of work by Newell and Simon (1972) and Hunt and his associates (Hunt, 1971, 1973; Hunt & Poltrock, 1974). Largely because of its generality, it is consistent with the analyses of information processing in more limited situations espoused by many authors.

Information processing includes the reception of physical signals, their classification as symbols bearing meaning with respect to previous experience, integration of identifiable units into a coherent whole, and derivation of new units of meaning by the interaction of currently presented and memorized information. Each of these stages can be identified even in intuitively simple acts of cognition, such as mental arithmetic.

We propose to discuss information processing abilities by considering the acts which must be executed at each stage. In order to do so, we find it useful to use an old and somewhat disreputable concept, *conscious thought*. People can reply sensibly to the question, "What are you thinking about right now?" Any information that reaches this level will be considered to be at the conscious thought level. In addition to such information, the human system will contain information that is currently being processed but has not yet reached the conscious level, and information that is being retained in memory but is not currently the focus of thought.

The system architecture of this model is shown in Figure 2. Preconscious information processing begins when a physical stimulus is sensed. The signal is assumed to be broadcast into long-term memory (LTM), where it may arouse some stored record, or *code*. The code may, in turn, be

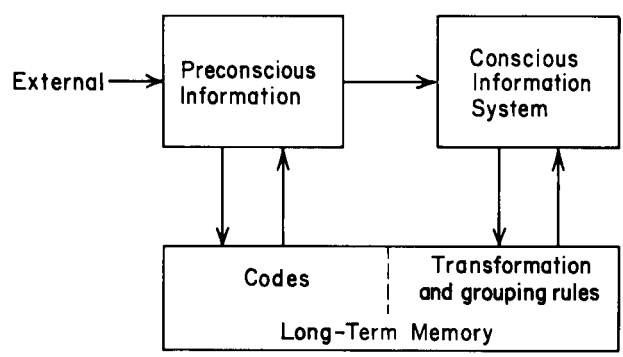

FIGURE 2 Model for temporal aspects of information processing.

rebroadcast into LTM, together with other recently identified codes, to arouse further, higher-level codes. At some point a code may be aroused which is at a sufficiently high level so that it enters conscious thought. Conscious thought is hypothesized to consist of two types of information. One is a "raw" collection of codes in short-term memory (STM), roughly equivalent to our span of immediately apprehended chunks of information. That is, STM is a temporally ordered record of those codes that have most recently arrived in conscious memory. In addition, we assume the existence of an intermediate-term memory (ITM) that holds a representation of the meaning of the current stimulus situation. The distinction is perhaps best grasped by considering the thought of a listener during a lecture. Short-term memory contains the listener's recollection of the speaker's last few words; ITM contains the listener's understanding of what the speech is about.[1]

Problem solving is seen as a sequence of transformations of information in the STM–ITM system, under the control of transformation rules ("productions," as introduced by Newell & Simon, 1972) that are stored in LTM. Some transformations may be grouped into sets (production systems), which, in effect, constitute memory of how to solve problems. We shall be more specific when we discuss problem solving itself.

D. Preconscious Information Processing

It has been well established that there are "short-term sensory stores" associated with the senses of vision and audition and possibly with the other senses. A sensory store can be defined an an internal representation of the physical characteristics of a stimulus, a representation that lasts for a very brief period of time following the cessation of the stimulus itself. The short-term visual store was first demonstrated by Sperling (1960), who showed that subjects had a visual representation of a digit matrix available to them for a fraction of a second after the presentation of the matrix. Sperling also demonstrated that this representation is purely visual, and has not been recoded to the semantic level; that is, the visual nature of the digits is available, but not their associated meanings. A similar sensory store has been shown to be associated with the sense of hearing (Crowder & Morton, 1969; Massaro, 1972), but its nature is much more controversial.

Although the sensory stores are believed to contain purely sensory information, several lines of research indicate that all stimuli that impinge on

[1] In our system, ITM is a repository of information for a period of minutes or even hours. Thus, it is operationally equivalent to *long-term memory* whenever that term is restricted to memory for events in an experiment which lasts for, say, an hour or two. Note, for instance, the use of LTM in the Atkinson–Shiffrin (1968) model.

the senses are automatically processed to a higher, nonsensory level (Corteen & Wood, 1972; Keele, 1973; Lewis, 1970). This processing occurs whether or not a stimulus is the present focus of attention. Furthermore, the preconscious processing of several stimuli appears to take place "in parallel," that is, simultaneously rather than serially. Lewis (1970), for example, presented pairs of words simultaneously to the subject's right and left ears. The subject was instructed to repeat the words presented to only one ear, the "attended" ear. Although subjects were not conscious of having heard the words which occurred in the other, "unattended," ear, and had no memory of them, these unattended words were apparently processed to the semantic level, since a semantic relationship between the attended and unattended words affected the subject's response time to the attended word.

Several factors determine whether preconscious information will enter consciousness. Information has a high probability of being attended and thereby entering consciousness if it enters the senses on a previously attended channel. However, if the code aroused by a stimulus presented on a previously unattended channel is of high priority, the stimulus itself may draw the subject's attention. For example, we know that one is sensitive to the sound of his or her own name, even when it occurs over a previously unattended airport public address system. In some situations, a cue on an attended channel may warn us to switch our attention, as when an aural cue signals a pilot to monitor a visual display.

Individuals may differ in their preconscious processes in at least three ways. The duration of the sensory store may differ. The speed at which stimuli are automatically coded may differ, and so may the level to which given classes of stimuli are coded. Finally, the speed and accuracy with which individuals can switch their attention in response to a high-priority signal may vary.

Individual differences in recoding and information-holding processes appear to be very large. In fact, experiments in this area are carefully designed to control interindividual variation. Consider Sperling's (1960) experiments, which established the existence of a visual iconic storage. Matrices of letters were presented for 50 msec. Subjects could name only a small proportion of the letters immediately after the display. However, if, immediately following the stimulus offset, a cue was given that indicated which row of the matrix the subject was to recall, he could name a much larger proportion of that row. The number of items that could be named decreased as the interval between stimulus offset and the cue increased. The experiment is interpreted as evidence that stimulus information is selectively transferred from a fading iconic storage into short-term (conscious) memory. In one of his experiments Sperling reported data from four subjects. The best performing subject had 85% of the stimulus information available in iconic storage after 150 msec, whereas the poorest

subject had only 60% of the information available at the same time. This finding is typical of those obtained in studies of information decay in the visual system, where it is normal to report data from individual subjects.

Differences are also evident in readout from short-term acoustic store. For example, Moore and Massaro (1973) asked subjects to recognize the loudness and/or quality of a test tone which was followed at a variable interval by a masking tone. Data for individual subjects indicate that the probability of correctly identifying both loudness and quality was twice as great for the best subject as for the worst.

Our theoretical formulation, in common with that of others (e.g., Craik & Lockhart, 1972) asserts that a main function of preconscious information processing is to associate a physical stimulus with a learned code. The "same–different" identification task developed by Posner and Keele (1967) provides a way to measure the recoding process directly. In the simplest form of this task the subject is presented with two alphabet characters and asked to respond "same" if they name the same letter. Pairs such as (A A) or (B B) could, of course, be identified as identical even if one did not know the letters they named since the patterns are physically identical. Pairs such as (A a) or (B b) can be identified as identical only by accessing the name code for the visual patterns. Posner and his associates argue that the difference in reaction time between *physical identity* cases and *name identity* cases can be taken as a measure of the time required for code access. In their experiments with unsystematically selected college students access time averages about 50–55 msec. Individual differences are quite large. We calculate that in one of their studies[5] over 85% of the variance in the name-access code effect was *not* associated with the "average" trend over subjects and thus must be due to either measurement error or individual differences.

Individual variation in the sorts of studies under discussion reflects differences in the rate at which the external world can be coded into a meaningful internal representation. Surely this variation must be an important aspect of cognition! Nevertheless, there have been few attempts to determine what effect individual variations in preconscious information processing have upon an individual's general cognitive ability. Work in our own laboratory indicates that there is some association between preconscious information processing and the processes measured by psychometric tests of verbal factors in intelligence. Subjects who score in the top quartile of the university student norms on a verbal intelligence test ("high verbals") are approximately 35% faster in code access time on the Posner–Keele same–different identification task than are students scoring in the lowest quartile ("low verbals") (Hunt, Frost, & Lunneborg, 1973;

[5] Using data from Posner, Boies, Eichelman, and Taylor (1969, Table 3). The calculation is for the "blank" and "mask" conditions. The calculation of ω^2 was done using the approximation formula provided by Lindman (1974, p. 54).

Hunt, Lunneborg, & Lewis, 1975). Similar, but less striking, findings have been obtained on tasks which may reflect decay of information in preconscious store as well as code access. High verbal students are able to read more information from store in the Sperling paradigm when the cue is given immediately after stimulus offset. They are also more accurate at reporting dichotically presented information when the cue is given after presentation of the information to be recalled, thus requiring that attention be switched to the appropriate channel. Hunt *et al.* (1975) presented digits and consonants dichotically, two digits and two consonants to each ear. They then cued the subject to recall either by categories (digits–consonants or the reverse) or by ear (right–left or the reverse). The ability to recall by categories was associated with psychometric verbal ability but the ability to recall by ear was not. These two results could be explained in either of two ways: the high verbals may have slower rates of decay or more rapid coding processes. We are inclined to favor the latter explanation. The interesting thing is that a person who is identified as being high verbal on the basis of test of vocabulary size and knowledge of grammar, tasks which apparently have nothing to do with preconscious information processing, is also likely to have faster access to the name code of a physical stimulus. This finding suggests that the gross psychometric classification also captures "verbal ability" at a more basic information-processing level.

We have noted that one of the important function of preconscious processing is to indicate when the observer should switch attention from one channel to another. Kahneman and his associates (Gopher & Kahneman, 1971; Kahneman, Ben-Ishai, & Lotan, 1973) had subjects monitor one of two dichotically presented lists. They were then cued either to change to monitoring the other ear, or to continue as before. Military combat aircraft pilots were more capable of reorienting their attention than were the pilots of military service aircraft, even though both groups exhibited a very high level of ability. The accuracy with which commercial truck drivers were able to identify the first few messages on the new channel was positively correlated with their safety records. One is hard put to imagine an individual difference more linked to survival in a transportation mad age.

E. Conscious Processing: STM and ITM

The existence of substantial individual differences in the retention, integration, and transformation of immediately presented information is well documented. Our earlier discussion of studies based on the Atkinson–Shiffrin model demonstrated substantial individual variation in an information-retention task, and also noted that this variation appears to be

associated with differences in information processing strategies rather than pure memory "capacity." Furthermore, an individual's use of strategies appears to have some correlation with his psychometrically defined intelligence. Similar conclusions have been reached by psychologists using other paradigms to study memory for information presented over a period of an hour or less. These studies suggest that differences in memory abilities are more often associated with either a general intelligence index or with the verbal component of an intelligence test, than with the quantitative ability component.[6]

Since STM plays such a central role in current theories of cognition, a measure of the size of an individual's STM would be of obvious importance. Miller (1956) has pointed out that STM size should be measured in terms of the number of "chunks," or integrated information units, that it can hold, rather than by some abstract quantitative measure of information. He has gone on to point out that the average STM size in chunks, measured *across* individuals, is remarkably constant for a wide range of materials. "Seven plus or minus two," Miller's estimate of the average STM size, has become one of the most widely quoted constants in psychology. The two most prominent individually administered intelligence tests, the Stanford–Binet (S–B) and Wechsler Adult Intelligence Scale (WAIS) both attempt to measure short-term memory span by asking the person examined to repeat back a sequence of spoken digits. This test has a correlation of about .60 with full-scale intelligence test scores (Wechsler, 1958). There is a temptation to regard this measure as a basic structural parameter of a person's cognition.

Estes (1974) has pointed out that the observed digit span depends upon how the subject organized chunks of information. It is not necessarily the case that one digit will equal one chunk. The mnemonist VP, previously referred to, improved his digit span over two sittings from a quite average 8 to a "super" 25. He did this by developing a sophisticated strategy for chunking successive digits into units, then remembering the units (Hunt & Love, 1972). Such a strategy sounds so simple that one wonders why everyone does not do it. In fact, it is quite difficult to form units from

[6] This result is not in conflict with those cited which indicated association of memory strategies and quantitative ability. In reaching this conclusion Hunt, Frost, and Lunneborg (1973) used a "pseudo-orthogonal" experimental design that balanced verbal ability and quantitative ability in the *experimental population*. In the population at large these abilities are generally positively correlated and hence studies using either composite intelligence or verbal ability scores alone could find positive correlations between such scores and any dependent variable, even though the balanced design used by Hunt and his associates would restrict the correlation to quantitative ability. Whether the balanced design or one that reflects the ability distribution in the population is important depends upon the generalizations one wishes to draw (Humphreys, personal communication).

rapidly presented information. Love (personal communication) attempted to teach university students to use VP's strategy to increase their digit span, but was unable to do so.

A second ostensibly simple memory task is supraspan memorization. This term refers to memorization tasks in which more material is presented than can be learned on a single repetition. In terms of our theoretical framework, the learning of long lists of material requires formation of structures in intermediate-term memory (ITM) that are capable of guiding a retrieval scheme for the material to be memorized.

Strategies for memorization are involved in supraspan memorization as much as in subspan memorization. Mandler (1967) has observed that any recall other than immediate memory span requires that the subject construct a scheme for retrieval of information, and that these schemes involve active organization of the data presented. Further research has shown that such retrieval schemes tend to be highly idiosyncratic and most certainly do *not* depend upon the simple learning of consecutive items in a list, in a paired-associates fashion (Mandler, Worden, & Graesser, 1974). Once again we see the effect of strategies on what is apparently a "rote-learning" task. Some psychometric theories of intelligence (notably Guilford, 1967; Jensen, 1969) distinguish between "rote memorization" and abilities involved in information transformation. This dichotomy is too simplistic. Tasks that were once thought to involve only passive memorization are accomplished by most college-educated, adult subjects (who, after all, loom large in the data base of psychology) by means of highly sophisticated strategies of information organization.

Experimental psychologists tend to regard strategies as nuisance variables. There seems to be an implicit assumption that we are interested in strategy-invariant truths, and thus should seek tasks in which strategies can be controlled. We think that this is a mistake. The ability to devise and use strategies for information handling should be seen as an important and stable individual variable. Some cross-cultural research suggests that the development of information processing strategies for a wide variety of tasks is related to amount of formal schooling (Cole *et al.*, 1971). Similar suggestions may be found in studies which have compared the performance of college students of different levels of education, for example, sophomores as opposed to seniors and graduate students (Hunt, Marin, & Stone, 1966, Chapter 7). We need to know whether formal education does indeed foster general information-processing strategies, and if it does, what in the educational process determines the generality with which students adopt a problem-solving approach to information acquisition and organization.

Certainly education and other social processes are important in establishing patterns of thought, yet there are also biological determinants to mental

capacity. This poses something of a paradox, since the individual differences we find are in strategies, and these are certainly not inherited or guided by a hormone. A possible resolution of this paradox is to locate a structural parameter which could indirectly determine the availability of various control processes. Speed of information processing is such a variable. Most immediate memory tasks are time paced, so that the execution of a strategy requires that relations between items in STM be established within a relatively short time. Complex strategies may require extensive processing which simply can not be done by a mentally slow individual within the time available. If this reasoning is correct, there should be demonstrable individual differences in rate of execution of basic steps in information processing, and these differences should be systematically related to other differences in cognition.

There is ample evidence for differences in information processing speed. Sternberg (1966) developed a method for measuring the time taken to scan an individual item in STM. First a small number of simple items (letters, digits) are presented in a *memory set* of size *n*. Then a *probe item* is presented, and the subject is asked if the probe item was a member of the memory set. The subject's reaction time is plotted as a function of the size of the memory set. The slope of this function is taken as a measure of the rate at which a single item can be scanned in STM.[7] If one varies experimental material, by conducting memory scan experiments using word, pictures, or nonsense syllables, there is a direct relationship between the STM scanning rate for a given type of material and the memory span for that material measured in experimenter-defined units (Cavanaugh, 1972). We, and others (e.g., Estes, 1972, 1974) have argued that STM span is determined by the formation of chunks of information. The chunks, in turn, provide the basic units that are fixed into (in our terms) ITM meaning-bearing structures. Cavanaugh's results argue that certain material can be scanned faster, and thus is more easily "chunked." Exactly the same argument applies if we consider the variation in scan rate over individuals, instead of over experimental material.

We should expect to find fairly large individual differences in speed of simple information processing in STM, and furthermore, we should expect these differences to relate to other important cognitive variables in some way. It is clear that there are large individual differences. Sternberg, in

[7] There is some controversy over the extent to which this assumption depends upon assuming a particular strategy for the manner in which items in STM are to be scanned. For a detailed discussion of the issues involved, see Townsend (1974). For the purposes of the argument we make here, this controversy is irrelevant. The reason for this is that the different measures proposed by various theorists as "the correct" estimate of STM scanning rate are all monotonically related, and our discussion depends only upon ordinal relations between scanning rates.

his research on scanning, finds such great individual variation that he typically reports data from individual subjects (Sternberg, 1966, 1970). In our laboratory we have found evidence that rapid STM scanning of verbal material is associated with high verbal intelligence, but the associations, while statistically significant, are surprisingly small (Hunt *et al.*, 1973, 1975). However, there are striking individual cases. The mnemonist VP had a scanning rate of 10 msec per letter, less than half the lowest time per letter reported in Cavanaugh's survey of the literature.

Our survey of information processing has concentrated on verbal material. However, it should not be concluded that the relevant cognitive variables are exclusively verbal, for it appears that the ability to handle visual information in short-term memory (for example, nonsense figures, pictures) is a somewhat different one (Kelley, 1964). Shepard and his

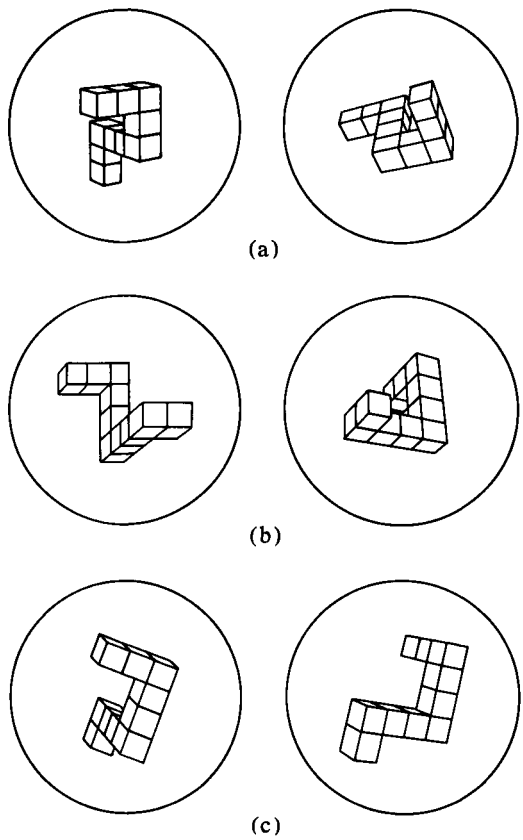

FIGURE 3 Examples of pairs of perspective line drawings: (a) identical pair differing only in rotation in picture plane; (b) identical pair differing by rotation in depth; (c) "different" pair which cannot be made identical by rotation. (From Shepard & Metzler, 1971, Fig. 1. Copyright © 1971 by the American Association for the Advancement of Science.)

colleagues have developed procedures for measuring the rate of manipulation of visual information in STM. A simple example of their procedure is shown in Figure 3. The subject is shown two pictures, and asked to determine whether they are of different objects, or of the same object observed from a different perspective. It has been found that the time required to answer this question is a linear function of the degree of rotation required to bring the figures to a common orientation (Shepard & Metzler, 1971). As is so often the case, wide individual differences are observed. It is of particular interest to note that when Shepard and his colleagues report individual data, the subjects appear to fall into two homogenous groups, "rapid" and "slow" subjects, with considerable differences between the group means (Cooper & Shepard, 1973; see also the individual data in a related visual task reported by Shepard & Feng, 1972). In this case we are fortunate to have rate of information processing measures anchored to other individual difference variables in two ways. Snyder (1973) and Yen (1973) have found correlations between rate of information processing in the Shepard and Metzler task and a number of psychometric tests of the ability to solve problems in spatial orientation. In addition, Yen (1973) and Bock and Kolakowski (1973) have analyzed the performance of related individuals on such tasks. Their data indicate that "visualization ability" follows a genetic pattern indicating the possibility of control by a single gene located on the X chromosome. Bock (1973), in a review of the research on this point, has predicted that further research will soon clarify the relationship between spatial ability and genetic makeup.

F. Problem Solving and Retrieval from Long-Term Memory

Although we maintain that memory tasks involve problem solving, problem-solving tasks require more than memory. Psychometricians have distinguished between *verbal comprehension* tasks, which involve knowledge of the language, and *reasoning* tasks, which also require the development of inferences from verbally stated facts. Lunneborg (1974) found that performance on a variety of "immediate information-processing" tasks, such as the Sternberg and Sperling paradigms, and the Posner–Keele same–different identification task, could be used to predict psychometric verbal comprehension scores, but could not be used to predict verbal reasoning scores. Similar illustrations have been made of the relation between visual information processing and quantitative problem solving. Hayes' (1973) experimental illustration of the role of imagery in solving even simple arithmetical tasks is illustrative. He showed that presenting an arithmetical task in such a way that "paper-and-pencil" operations could be visualized facilitated solution. At a much higher level of mathematics, creative mathematicians

frequently report the use of imagery to establish models that suggest solutions to problems (Polya, 1954). Surely, no one would claim that there was no more to mathematics than visualizing!

In order to be more specific, we shall fit "problem solving" and "retrieval of information from long-term memory" into our model using the Newell and Simon (1972) concept of *productions* and *production systems*. We have already introduced the idea of a production at an intuitive level. A production is a rule of the form

$$A \rightarrow B;\ C, \tag{1}$$

where A and B are classes of states of conscious memory and C is an overt action. By *classes of states of conscious memory* we mean that A and B specify general expressions which can be instantiated in a specific case, in much the same way that the algebraic expression $X + Y$ can be instantiated by the expression $2 + 4$. A production is activated when an A form is recognized in conscious memory. The A form is then replaced by the B form. As an example, we can express the distributive rule of algebra by

$$A \cdot (B + C) = A \cdot B + A \cdot C, \tag{2}$$

and the rule for changing an English language passive sentence to its active form by

(3) (NOUN 1)(PASSIVE VERB FORM)(NOUN 2) →
(NOUN 2)(ACTIVE VERB FORM)(NOUN 1).

Newell and Simon's basic assumption is that in problem solving the successive states of conscious memory can be thought of as the result of activation of the appropriate productions. As we have pointed out, the simplest form of a production is a recoding, as in the change from the auditory code "DAWG" to the internal semantic code for the animal. Problem solving is seen as different from recoding only in that it involves successive activation of a series of productions.

Newell and Simon regard LTM as a repository for ordered sets of productions, called *production systems*. These are effectively programs that, when activated, can produce quite complicated changes in the contents of conscious memory, without the intervention of outside stimulation. Newell and Simon and their collaborators have written production systems capable of reproducing (*simulating*) the behavior of individuals engaged in problem solving situations as difficult as chess playing, the solution of symbolic logic problems, or the solution of verbally stated "puzzles" of the type found in articles on mathematical games. As one might expect, it is usually necessary to adjust a production system to each individual whose behavior is to be simulated, so the Newell and Simon approach provides a very fine grain analysis of individual differences. On the other

hand, this analysis is so fine that it is difficult to systematize it in any way beyond stating the fact that individual differences have, indeed, been observed.

Two general principles about problem solving have been developed from the Newell and Simon analysis. The first is the idea of a *representation.* In order to match productions to internal conscious representations of an exterior problem, there must be some coding for the external problem. What coding is chosen will determine what productions will be activated initially, and therefore what the course of problem solving will be. To illustrate, consider the case of Euclidean geometry. Strictly speaking, this is a system in formal reasoning in which inferred sentences are to be defined from the sentences in the premise, using only precisely defined rules of inference. In practice virtually everyone draws a diagram and then reasons about the angles and lines in the diagram. This works because there is a close correspondence between the properties of lines and angles and the definitions of plane geometry. Thus, we can go back and forth between problem solving in the diagram and in the formal system.

The second important principle is that of internal data management. Production systems are, in effect, programs, and they may have to generate internal codes, markers, and even call other programs as subprograms. Like any program, a production system must carry out these actions within the structural confines of the memory available to it. If the internal codes overflow the internal memory structures, then the production system (that is, the problem solver) will literally lose its way, and be unable to generate a systematic attack on the problem. This would clearly be the case for a production system realized as a computer program. Do humans behave in the same way? Probably they do. Malin (1973; see also Greeno, 1973) has shown that simple algebra problems are easy or hard for college students in proportion to the number of opportunities they offer for production systems to enter profitless sequences of transformations or to generate excessive internal codes.

Granted that these are reasonable concepts for computer programs, can they be applied to the study of human individual differences? To show that they can, we should show that human problem solving does depend upon finding representations and upon developing an orderly procedure for transforming problem representations. Surprisingly, there is excellent evidence for this dependence which antedates the development of the "information processing" approach to thought.

Some years before the computer analogy became popular in psychology, Bloom and Broder (1950) contrasted the problem solving performance of good and poor University of Chicago students. Since the University of Chicago was at that time one of the most intellectually oriented colleges in the country, Bloom and Broder's subjects were all of relatively high

TABLE 2
Typical Problem Used by Bloom and Broder[a]

Statements:

Any action that impedes the war effort of the United States should be made illegal.

All strikes impede the war effort of the United States.

Which of these conclusions follow?

A. All strikes should be made illegal.
B. Some restrictions should be placed on the right to strike, but it would be unwise to make them illegal.
C. Some strikes should be made illegal.
D. Unjustifiable strikes should be made illegal.
E. None of the foregoing conclusions follows.

[a]From Bloom and Broder (1950).

mental ability. This makes their results still more interesting. Bloom and Broder asked students to solve "work problems" similar to those found on college and graduate school entrance examinations. An example is shown in Table 2. The students were asked to "talk aloud" while solving these problems, and then were asked probing questions about how they attacked the problem. The summary statements of Bloom and Broder's interviews read remarkably like today's computer science statements of problem solving:

1. The better students were able to isolate the key elements of a problem and the relations between these elements. They established representations (our term) that retained and highlighted such information. Poorer students did not focus on key elements of the problem and often used a representation which did not contain important parts of the problem statement.
2. Poorer students often did not read the instructions carefully, and thus did not understand the requirements of an answer. In computer science terms (and in Newell and Simon's), they did not have a clear understanding of the target state toward which they were asked to transform the initial information.
3. Good students investigated the implications of each statement in a systematic way. The actual way used varied greatly between individuals. Poor students jumped from one transformation of information to another, without formulating their plan of attack.
4. Poor students sometimes reacted emotionally to irrelevant aspects of the problem statement. There is no analogy to such behavior in the

Newell and Simon work, although there is nothing inherent in the productions system approach which would prevent modeling this aspect of human thought.

Bloom and Broder also observed that, regardless of ability, individuals had their own characteristic ways of looking at problems. Some people preferred to visualize a problem, others to represent it symbolically, and still others consistently sought analogies between the present problem and some previously solved one. The first two classes of representation represent a link in a chain of similar observations tracing back at least to Galton (1883), who speculated that people fall into relatively discrete classes of "verbalizers" and "visualizers," with respect to their characteristic mental representation of problems. This speculation has been investigated many times. Paivio (1971), after an extensive review of the evidence, concluded that the abilities to deal with problems via verbal or visual codings appear to be independently distributed. Most people have a limited capacity to deal with problems in either way, while a few individuals may rely on one coding to the exclusion of the other. [For contrasting pictures of two different individuals, see Luria's (1968) account of an almost exclusively visual coder, and Hunt and Love's (1972) description of a verbal coder.] Because such extreme individuals are unusual, they tend to impress themselves on our memory, and may lead us to the erroneous conclusion that there are discrete types.

The title of this section refers to both problem solving and long term memory, and our final remarks will attempt to connect these two topics. We tend to think of LTM as a repository of a record of past events. Perhaps it would be more accurate to think of it as a repository of knowledge about how to do things. Norman and Rumelhart (1975), in presenting a detailed hypothesis about the structure of LTM, stress that we do not remember what happened—for example, the steps we took in scrambling an egg for the first time—but rather we remember a procedure for doing things. Viewed in this way, Newell and Simon's production systems are simply notational conventions for describing procedures. We are still faced with the question of how the procedures are aroused. The only sensible answer we can give is that the encoding of the procedure *at the time when it is stored* is a key step. Thus, it is not enough to teach someone a strategy for memorizing, one must also ensure that the subject knows when to use it. Any individual differences in handling immediately presented information will determine the amount of encoding that will be done when a procedure is learned. Thus, the effect of a small difference in the encoding of a problem solving procedure will be repeated every time the person encounters a situation in which the encoded procedure might be a useful approach to problem solving. We suggest that this indirect effect on

problem solving may be one of the most important effects of differences in human information processing, and that explicating this effect may be one of the most important reasons for studying these differences.

II. CONCLUSION

We began with a critique of the psychometric approach to the study of individual differences in cognition, pointing out that although the techniques developed by psychometricians do provide a reliable means of classifying people along various ability dimensions, they do not give explanations of the mental processes which underly the psychometrically defined abilities. Neither do they give enough information as to why some people are superior to others in these processes. They do not help us find the mechanisms which connect environmental and physiological variables to abilities. We proposed that a satisfactory theory of individual differences should be based on concepts which are rooted in a theory of cognition.

While we were reviewing the literature for this presentation, at times we were tempted to conclude that the number of ways in which people differ is so vast that there is no hope of finding order in the chaos. Perhaps the assumption of the factor analysts that mental performance is based on a limited number of ability dimensions is a useful fiction, since it allows the investigators to keep their sanity. We are now more optimistic. There is hope for progress in attacking the problem of individual differences using the models provided by modern theories of information processing.

Reexamination of some classic studies of preconscious information processing suggested that there are large differences in readout from iconic and echoic store, in recoding from the sensory to the name level, and in ability to switch attention on cue. More recent studies, including those of our own associates, indicate that these differences are associated in various degrees with differences on the psychometric variable of verbal ability.

Newell and Simon's concept of problem solving as the activation of a set of "production systems"—rules for the transformation of one set of symbols into another—has provided a promising basis for the study of the individual's problem solving ability. We noted that the characteristic differences between good and poor problem solvers isolated by Bloom and Broder (1950) could be reinterpreted in terms of Newell and Simon's production system approach.

The technique commonly used by the psychometrician to study the structure of abilities has been to administer a battery of diverse tests to a set of subjects and to attempt to extract information from the correlations between those tests. Those interested in developing a theory of individual differences that is related to the conceptual structure of experimental psychology have turned to several other techniques. One such technique is

to look at the correlation in performance on a small number of simple tasks all of which have been analyzed in terms of the information processes they involve. Another is to divide a group of subjects according to their ability on a complex task, and then compare the performance of the subgroups on a set of simpler tasks. Studies of nontypical individuals such as the retarded or the mnemonist have also added to our understanding of human abilities.

But we believe that the most important contributions will come from experimental psychologists who take the time to look at their data not only in terms of group means, but in terms of individual subjects. In the past, many psychologists have seen individual differences as sources of error variance to be eliminated through the use of efficient experimental designs. The complexity of the causal chain between stimulus and response has forced them to turn their attention away from individual data to group means. There is a real danger that the laws revealed by these mean scores are not true of any individual subject, but are the result of averaging two or more distinct, nonmodal patterns of behavior.

The ultimate goal of psychology is to understand and predict the behavior of the individual subject, not of the composite college sophomore. Thus, it seems reasonable that after we feed our data to the computer to determine group trends we go back and see to what extent these trends are typical of individual subjects, and that where they are not, we seek explanations.

ACKNOWLEDGMENTS

The preparation of this paper was supported by the National Institute of Mental Health, grant MH 21795, to the University of Washington.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. Spence & J. Spence (Eds.), *The psychology of learning and motivation*. Vol. II. New York: Academic Press, 1968.

Belmont, J. M., & Butterfield, E. C. The relations of short-term memory to development and intelligence. In L. P. Lipsitt & H. W. Reese (Eds.), *Advances in child development and behavior*. Vol. 4. New York: Academic Press, 1969.

Bloom, B., & Broder, L. *The problem solving processes of college students*. Chicago: University of Chicago Press, 1950.

Bock, R. D. Word and image: Source of the verbal and spatial factors in mental test scores. *Psychometrika*, 1973, **38**, 437–458.

Bock, R. D., & Kolakowski, D. Further evidence of sex-linked major gene influence on human spatial visualizing ability. *American Journal of Human Genetics*, 1973, **25**, 1–14.

Brown, A., Campione, J., Bray, N., & Wilcox, B. Keeping track of changing variables: Effect of rehearsal training and rehearsal prevention in normal and retarded adolescents. *Journal of Experimental Psychology*, 1973, **101**, 123–131.

Brown, A., Campione, J. C., & Murphy, M. Keeping track of changing variables: Long term retention of a trained rehearsal strategy by retarded adolescents. *American Journal of Mental Deficiency*, 1974, **78**, 446–453.

Cattell, R. B. *Abilities: Their structure, growth, and action*. Boston: Houghton-Mifflin, 1971.

Cavanaugh, J. P. Relation between the immediate memory span and the memory search rate. *Psychological Review*, 1972, **79**, 525–530.

Cole, M., Gay, J., Glick, J., & Sharp, D. W. *The cultural context of learning and thinking*. New York: Basic Books, 1971.

Collins, A., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 240–247.

Cooper, L., & Shepard, R. Chronometric studies of the rotation of mental images. In W. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.

Corteen, R. S., & Wood, B. Autonomic responses to shock-associated words in an unattended channel. *Journal of Experimental Psychology*, 1972, **94**, 308–313.

Craik, F. I. M., & Lockhart, R. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 671–684.

Crowder, R. G., & Morton, J. Precategorical acoustic storage (PAS). *Perception and Psychophysics*, 1969, **5**, 365–373.

Edwards, A. L. *The measurement of personality traits by scales and inventories*. New York: Holt, Rinehart, & Winston, 1970.

Estes, W. K. An associative base for coding and organization in memory. In A. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington: Winston, 1972.

Estes, W. K. Learning theory and intelligence. *American Psychologist*, 1974, **29**, 740–749.

Galton, F. *Inquiry into human faculty and its development*. New York and London: Macmillan, 1883.

Gopher, D., & Kahneman, D. Individual differences in attention and the prediction of flight criteria. *Perceptual and Motor Skills*, 1971, **33**, 1335–1342.

Greeno, J. The structure of memory and the process of solving problems. In R. Solso (Ed.), *Contemporary issues in cognitive psychology*. Washington, D.C.: Winston, 1973.

Guilford, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.

Guilford, J. P., & Hoepfner, R. *The analysis of intelligence*. New York: McGraw-Hill, 1971.

Harman, H. *Modern factor analysis*. (2nd ed.) Chicago: University of Chicago Press, 1967.

Hayes, J. R. On the function of visual imagery in elementary mathematics. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.

Hunt, E. What kind of computer is man? *Cognitive Psychology*, 1971, **2**, 57–98.

Hunt, E. The memory we must have. In R. Shank & K. Colby (Eds.), *Computer models of thought and language*. San Francisco: Freeman, 1973.

Hunt, E., Frost, N., & Lunneborg, C. Individual differences in cognition: A new approach to intelligence. In C. Bower (Ed.), *Advances in learning and motivation*. Vol. 7. New York: Academic Press, 1973.

Hunt, E., & Love, L. T. How good can memory be? In A. Melton & E. Margin (Eds.), *Coding processes in human memory*, 1972. Washington, D.C.: Winston, 1972.

Hunt, E., Lunneborg, C., & Lewis, J. What does it mean to be high verbal? *Cognitive Psychology*, 1975 (in press).

Hunt, E., Marin, J., & Stone, P. J. *Experiments in induction.* New York: Academic Press, 1966.

Hunt, E., & Poltrock, S. The mechanics of thought. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1974.

Jensen, A. How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 1969, **39**, 1–23.

Kahneman, D., Ben-Ishai, R., & Lotan, M. Relation of a test of attention to road accidents. *Journal of Applied Psychology*, 1973, **58**, 113–115.

Keele, S. W. *Attention and human performance.* Pacific Palisades, California: Goodyear, 1973.

Kelley, H. P. Memory abilities: A factor analysis. *Psychometric Monographs*, 1964, No. 11, 1–53.

Lanvin, D. E. *The prediction of academic performance.* New York: Russell Sage Foundation, 1965.

Lesser, G., Fifer, G., & Clark, D. Mental abilities of children from different social-class and cultural groups. *Monograph of Society for Research in Child Development*, 1965 (Whole No. 105).

Lewis, J. Semantic processing of unattended messages using dichotic listening. *Journal of Experimental Psychology*, 1970, **85**, 225–228.

Lindman, H. R. *Analysis of variance in complex experimental designs.* San Francisco: Freeman, 1974.

Love, L. T. Location and analysis of exceptional memory. Unpublished master's thesis. Psychology Department, University of Washington, 1973.

Lunneborg, C. E. Address given at American Psychological Association (Division 3) symposium on individual differences in cognition, New Orleans, 1974.

Luria, A. *The mind of a mnemonist.* New York: Basic Books, 1968.

Malin, J. E. T. An analysis of strategies for solving certain substitution problems. University of Michigan, Human Performance Center, Technical Report No. 40, April, 1973.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation.* Vol. I. New York: Academic Press, 1967.

Mandler, G., Worden, P., & Graesser, A. C., II Subjective disorganization: Search for the locus of list organization. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 220–235.

Massaro, D. W. Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, 1972, **79**, 124–145.

Matarazzo, J. D. *Wechsler's measurement and appraisal of adult intelligence.* (5th ed.) Baltimore: Williams & Wilkins, 1972.

Miller, G. The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81–97.

Millward, R. B., Rice, G., & Corbett, A. Category production measures and verification times. In A. Kennedy & A. Wilkes (Eds.), *Studies in long term memory.* New York: Wiley, 1975, in press.

Moore, J., & Massaro, D. Attention and processing capacity in auditory recognition. *Journal of Experimental Psychology*, 1973, **99**, 49–54.

Nelson, T. O. Savings and forgetting from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 568–576.

Newell, A., & Simon, H. *Human problem solving.* Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Norman, D. A., & Rumelhart, D. E. *Explorations in cognition.* San Francisco: Freeman, 1975.

Paivio, A. *Imagery and verbal processes.* New York: Holt, Rinehart & Winston, 1971.

Polya, G. *Mathematics and plausible reasoning.* Princeton, New Jersey: Princeton University Press, 1954.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. Retention of visual and name codes of single letters. *Journal of Experimental Psychology,* 1969, **79,** No. 1, Pt. 2, 1–16.

Posner, M. I., & Keele, S. W. Decay of visual information from a single letter. *Science,* 1967, **158,** 137–139.

Shepard, R. N., & Feng, C. A chronometric study of mental paper folding. *Cognitive Psychology,* 1972, **3,** 228–243.

Shepard, R. N., & Metzler, J. Mental rotation of three dimensional objects. *Science,* 1971, **171,** 701–703.

Snyder, C. Unpublished doctoral dissertation. University of Oregon, 1973.

Spearman, C. *Abilities of man.* London: Methuen, 1927.

Sperling, G. The information available in brief visual presentations. *Psychological Monographs,* 1960, **74**(Whole no. 498).

Sternberg, S. High speed scanning in human memory. *Science,* 1966, **153,** 652–654.

Sternberg, S. Memory·scanning mental processes revealed by reaction time experiments. In J. Antrobus (Ed.), *Cognition and affect.* Boston: Little, Brown, 1970.

Thurstone, L. L. Primary mental abilities. *Psychometric Monograph,* No. 1, 1938.

Townsend, J. T. Issues and models concerning a finite number of inputs. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1974.

Tyler, L. E. *Individual differences: Abilities and motivational directions.* New York: Appleton-Century-Crofts, 1974.

Tyler, L. E. The intelligence we test—an evolving concept. In L. Resnick (Ed.), *The nature of intelligence.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, in press.

Vernon, P. E. *The structure of human abilities.* London: Methven, 1961.

Wechsler, D. *The measurement and appraisal of adult intelligence.* (4th ed.) Baltimore: Williams & Wilkins, 1958.

Yen, W. Sex-linked major gene influences on human spatial performance. Unpublished doctoral dissertation, University of California, Berkely, 1973.

4

Comparative Psychology and Human Cognition

Douglas Medin
Michael Cole

Rockefeller University

One of the strange aspects of writing an article advocating comparative research on human cognition in the latter quarter of the twentieth century is that the idea is at least a century old. Darwin's famous study of the expression of emotions in man was but one of many essays in comparative psychology in the last century. In the current century there have been several attempts to incorporate comparative research into the main body of psychological theory (Werner, 1948).

Generally, however, these attempts were not particularly influential, mainly, in our opinion, because neither cognitive theories nor comparative research strategies had evolved to the point where there was a strong basis for positive interaction between the two.

In this presentation we would like to consider the theoretical and practical advantages to be gained from the inclusion of a wide range of subject populations and behaviors in a general psychology of human cognition. Our goal is to provide a framework for evaluating and integrating data gathered from diverse subject populations. We hope to demonstrate that this framework can be useful to the comparative researcher interested in the basis of group differences and can be relevant to understanding cognitive processes independent of any special interest in comparative enterprises. The domain of comparative research in this article is intended to include at least phylogenetic, developmental, and cross-cultural comparisons because we believe that there is a logic to the comparative approach

that is basically similar, regardless of whether we are looking at comparisons across cultures, across species, or across ages within a species. Further, we maintain that studies of special groups, whether they are particular species, ages, or cultural groups, can be more enlightening if the research is appropriately guided by theoretical considerations.

The most common rationale for comparative, cognitive research, which we shall call the search for historical antecedents, dates at least back to Darwin. Phylogenetically, ontogenetically, or culturally distinct groups are treated as providing "antecedent" conditions, which the psychologist treats as the "causes" of the behavior in the reference groups for which his theory has been developed. (Does language acquisition cause a change in memory? Does a developed frontal lobe cause greater flexibility in problem solving?)

The logical problems inherent in such antecedent–consequent research have been extremely difficult to resolve. A central paradox resides in the fact that the initial differences which motivate our interest in the comparative enterprise in the first place, simultaneously violate the *ceteris paribus* assumption upon which classical experimental cognitive research depends. These issues have been discussed by many competent investigators (Bitterman, 1965; Campbell, 1961).

Although we think the issues involved in such applications of the comparative method are of scientific importance, we shall concentrate on a second rational for comparative research, which (following Lewin, 1935) we will call the search for "systematic" explanations—we will seek to elucidate the mechanisms operating within an individual at the time of observation that give rise to the performance we observe. We are narrowing the topic in this way primarily because we believe that clearer answers to the "causal-antecedent" questions of comparative research require a clearer conception of how cognitive processes operate *within* individuals comprising the groups being compared

I. SOME COMPARATIVE RESEARCH STRATEGIES

A. The Method of Equation

One common comparative research strategy has been to select a single task and to compare the performance of two or more groups on "equivalent" varieties of the task. In the early part of this century it was popular to make both phylogenetic and ontogenetic comparisons using a variety of simplified problem-solving tasks; the Hampton Court maze, Thorndike's puzzle box, Hunter's obstruction box, and Köhler's box-stacking problem all come to mind as examples that have been popular in comparative work

at one time or another. When these comparisons are made using what we term "the method of equation," the experimenter makes every effort to equate all relevant factors. For example, when comparing the performance of two species on some learning task, the attractiveness of the reward, the response requirements, the amount of deprivation, and the saliency of the stimuli are just some of the factors that must be equated if the comparisons are to be "fair."

It is certainly important in any study to control for unwanted sources of variability and a set of experiments using the method of equation can often produce a meaningful picture with respect to some question. However, frequently there is no general solution to "making all things equal" when we are working with naturally occurring groups that differ in various ways. There is also increasing recognition that, even if all other things were equal, there is no logical necessity that equal performances should imply identical underlying mechanisms (Werner, 1957).

B. The Method of Functional Relationships

Criticism of simple "equation" comparisons led to the development of many different strategies for making intergroup comparisons more informative. One of the most popular approaches has been to seek equivalence of functional relations. Rather than seek to establish control over all potentially significant sources of variation in a single experiment and look at differences with respect to a single parameter, one investigates potentially relevant sources of variation in a systematic fashion. Similarities and differences in the influence of relevant factors are the focus of interest. This is the approach used by Skinner (1959) in his argument that the laws of reinforcement apply across species; the cumulative curves reflecting the influences of periodic reinforcement on bar-pressing rate show a similar pattern for the pigeon, rat, and monkey.

When such functional relations differ among species, the hypothesis of different learning mechanisms gains credence. Thus, Bitterman (1965) reports a partial reinforcement-extinction effect in rats (that is, the animals respond longer following the removal of partially applied than 100% reward), but the opposite relation in fish. If this pattern of results does not interact with such factors as trial spacing, reward magnitude, or deprivation, it might be presumed that a fundamental difference in learning mechanisms in the two species has been demonstrated.

The method of functional equivalence clearly has an advantage over the method of equation. The analysis of functional relations leads us to assess performance in a variety of tasks, and, when successful, this method can rule out trivial explanations of group differences. But the search for functional equivalence also has its problems. For one thing, it does not provide

a guide for deciding how to proceed if we obtain a groups by conditions interaction. What can we conclude if the direction of the partial reinforcement effect varies differently as a function of trial spacing or deprivation levels in the rat and the fish? An equally troublesome problem arises when the results do not lead to a rejection of the null hypothesis and strong inferences are drawn from the absence of an effect. If we fail to get an interaction with our chosen trial-spacing and deprivation conditions, is it because no such interactions exist, or because we have simply failed to find the proper amount of trial spacing or severity of deprivation which would allow an interaction to manifest itself?

C. Process Orientations

The shortcomings of the two approaches we have just sketched reduce to a single problem: when we observe differences in behavior between two groups, we do not know what inferences are warranted concerning the mechanisms that give rise to the performance of individuals within each of the groups. In recent years there has been a growing consensus that cognitive mechanisms can be viewed more appropriately as processes than as stable, fixed quantities or discrete entities. Since processes might be brought to bear in a number of situations producing a variety of performances, it seems more pertinent to ask what kind of mechanism is at work in some population or situation than to ask how much capacity or ability subjects in that population possess (see Stevenson, 1972). Since the kinds of processes or abilities that many students of developmental and abnormal psychology (to name but two areas) want to make inferences about are hypothetical constructs, such as memory, it is almost a necessity that comparisons be made within a theoretical context. In short, some kind of theory is needed to guide our interpretation and to order the phenomena we are studying.

II. DEVELOPMENTS IN THEORY CONSTRUCTION

Much of the theoretical effort in the 1930s and 1940s was oriented toward extremely broad descriptions of behavior, perhaps the most influential being Clark Hull's (1943) system. Hull is best known for developing general laws of behavior that were presumed to characterize basic processes such as discrimination learning, trial-and-error learning, primary and secondary motivation. Hull (1945) was also interested in the problem of individual and species differences. If the primary laws of behavior are applicable to all mammals, the proper place to express species or

individual differences is in the parameters of the equations that express the general laws.

Estes (1970), in reviewing the application of Hullian theory to mental retardation research, suggests some reasons for its relative lack of productivity in that comparative effort. Discussing the problems of inference from a demonstration that Hull's habit-strength equation could fit the learning curves of normal and retarded subjects on a serial learning task, Estes points out that the single experiment is insufficient to isolate the source of the observed differences. While plausible differences in learning rate parameters are obtained, Estes suggests that a series of experiments varying factors that enter into the theory would be necessary to assure us that factors other than those reflected in the observed parameter differences are not the real sources of group differences. Few such studies have been forthcoming, at least partially because the parameters in Hull's system, such as variability in oscillatory inhibition, did not have the psychological content that would motivate comparative research interest.

The 1950s and 1960s were characterized by more limited goals in theory construction as the global theories or systems of the 1940s were superceded by models attempting to describe behavior in single experimental situations. While this line of work has elucidated aspects of behavior in narrowly defined situations that may have a more general significance, there has been little for the comparative researcher to garner from such work, precisely because of uncertainty as to what extent the behavior under observation was "situation specific."

Recently, we have witnessed the development of models that have proved capable of accounting for data from a number of different experimental paradigms with the added virtue that the parameters of the models appear to have a more distinct psychological content.

Collectively, the maturation of comparative research tactics and developments in theory construction have set the stage, we believe, for a mutually beneficial interplay between these two domains.

In this chapter we try to show that a well-specified theory can help to guide comparative research when we encounter "unexpected" results and that such a theory can suggest appropriate procedures for bringing group differences in a particular process into better focus. Briefly put, a process orientation applied in a theoretical framework can serve to sort out what observed differences are differences in.

We also try to elucidate reciprocal benefits for cognitive theory. A fundamental purpose of any theory is to relate disparate observations. A corollary of this purpose is the general scientific goal of maximizing both the generality and the parsimony of the theory. The desire for generality leads to a search for systematic variation to provide new tests of the theory.

Comparative research is one source of variation that can lead to greater refinement of the theory, or betray it as inadequate or incomplete.

In the following sections of this presentation we discuss several research topics, in each case attempting to point out developments which set the stage for comparative studies. We then try to derive some lessons concerning comparative and cognitive research from these studies and, where possible, make recommendations for future work.

III. SELECTIVE LEARNING

At least since the early 1940s, the study of selective mechanisms in discrimination learning has been carried out with a broad range of subject populations, and in many cases interpopulation comparisons have been an integral part of the research and theory.

One of the best known studies in this genre was conducted by Kuenne (1946), who worked with four- and eight-year-old children, but included a discussion of work with monkeys and rats as part of her theoretical rationale. Kuenne was concerned with the nature of the effective stimulus for children of different ages, and came to the well-known conclusion that young children, like some animals, respond to the absolute properties of stimuli whereas older children, using language as a mediator, respond to stimulus relations. While several later analyses have brought Kuenne's conclusions into question (the monograph by Bryant, 1974, is especially cogent), the tradition of including age, species, and other comparisons as part of discrimination learning theory has remained strong.

A central issue in comparative studies of selective learning concerns whether age and specific differences reflect qualitative differences in the learning process or whether the learning processes are fundamentally the same, but differ in their quantitative manifestations. Such a question does not admit to easy answers, but to the extent that theories have been able to provide a focus for this question, our understanding of the nature of selective learning and group differences in selective learning has been enhanced.

One of the early and comprehensively comparative approaches to a general theory of selective learning was developed by Kendler and Kendler (1962). The major experimental vehicle for their work has been the study of discrimination transfer, especially the kind of problem that has been labelled "reversal and nonreversal shifts."

In almost all problems of this type, the stimuli differ simultaneously on two dimensions (for example, color and form). The subject is rewarded for responses to one dimension (for example, red is correct, green is incorrect), while the other dimension is irrelevant. That is, a subject may be

presented a red triangle and a green circle on some trials and a red circle and a green triangle on other trials, and the reward is always associated with the red forms.

Following the achievement of criterion performance on the initial task, the problem is changed by altering the stimulus-outcome relations. For a reversal shift, the previously correct cue (red) becomes incorrect, and the previously incorrect cue (green) becomes correct. For a nonreversal shift a different dimension becomes relevant. For example, "triangle" may become correct and the subject must choose the red triangle over the green circle and also the green triangle over the red circle. The Kendlers observed that younger children and slower learners solved nonreversal shifts faster than reversal shifts, whereas the opposite was true for older children and faster learners.

According to the Kendlers, this difference in the relative rates of learning these two shifts reflects a qualitative difference in the underlying learning process. They propose that for animals and young children, learning involves a direct link between absolute properties of the stimuli and the appropriate response tendencies by the simple stimulus–response association principles espoused by Spence (1937) and Hull (1943). However, in the case of older children and adults, the Kendlers suggest that learning is modulated by a mediational process (possibly, but not necessarily, verbal) that allows learning and transfer to occur at the level of dimensions, and not just at the level of specific cues. For animals and young children, nonreversal shifts mean a change in only one of two stimulus–response connections (for example, when triangle becomes correct after red has been correct the subject must continue to choose the red triangle and learn to choose the blue triangle). Therefore, these shifts occur more rapidly than reversal shifts, which require change in two stimulus–response connections.

The interpretation of the theory is more complex for older children. They need to change the mediator as well as the specific stimulus–response relations for the nonreversal shift. For reversal shifts they can maintain the same mediator, switching only the particular stimulus response relations, which renders the reversal shift relatively easy in comparison to the nonreversal shift.

The Kendlers have applied their theory to a broad range of problem-solving phenomena, and it is not surprising that it has come in for a good deal of criticism on both empirical and theoretical grounds. Some researchers have failed to observe the basic age and species-related interactions predicted by the theory. Others have offered alternative theories that seem to account for more details of the data (see Zeaman & House, 1974, for a recent summary of the implications of the Kendlers' theory as well as their own).

One avenue to improving inferences concerning the comparative data on reversal and nonreversal shift performance is to consider aspects of the data more detailed than relative error frequencies. This is the approach taken by Tighe, Glick, and Cole (1971), who separated their analysis of performance on a nonreversal shift in terms of its two subproblems, one for which the reward conditions were changed and one for which the reward conditions were unchanged during transfer. For example, in the shift from red to triangle as correct, the unchanged subproblem involves comparison of a red triangle and a green circle and the changed subproblem involves a green triangle and a red circle. These investigators noted that in cases in which young children learned the nonreversal problem quickly, they made virtually no errors on the unchanged subproblem during transfer, almost as if the changed and unchanged subproblems were being treated as completely independent discriminations. Furthermore, performance on the changed subproblem in the nonreversal shift approximated performance on each of the reversal subproblems (both of which were changed).

Older children, on the other hand, manifested severe disruption on the unchanged subproblem, and their performance on the changed subproblem was inferior to that on subproblems of the reversal shift. In some cases (Cole, 1973) this subproblem analysis shows very large differences in performance that neither the total error nor the relative error rates reveal. These results raise the possibility that the main age differences are not differences in the presence or absence of mediation but rather are differences in whether or not the two subproblems are perceived as related or independent.

This subproblem analysis suggests that under some circumstances the age or species differences in reversal and nonreversal shift performance imply a qualitative change in the nature of the basic learning mechanisms. This conclusion is supported most strongly in cases in which there is errorless performance on unchanged subproblems while changed subproblems are being relearned.

A theory which challenges the existence of qualitatively different learning mechanisms associated with age or phylogenetic status is the attention theory of Zeaman and House (1963, 1974). Their work was originally applied to the discrimination performance of retarded children, but has become a very influential theory of discrimination learning in normal children. Recent formulations of the theory account for performance in a variety of tasks, including discrimination transfer.

Zeaman and House were motivated by an examination of the learning curves for individual retardates on problems differing in difficulty. Invariably it seemed that the learning curve took the form of a flat, chance-level performance followed by a relative abrupt rise to problem solution.

Performance differences between the various subgroups and levels of problem difficulty seemed to occur in the duration of the flat portion of the learning curve and not in the slope of the curve once it departed from this chance level.

To explain these data, Zeaman and House offered a two-stage analysis of discrimination learning. The main idea was that solving a discrimination problem requires subjects to learn to attend to the relevant stimulus dimension and then to learn to choose the correct cue on the relevant dimension.

Dickerson, Novik, and Gould (1972), working within the Zeaman–House framework, suggested that compared with older children, younger children have a slower rate of changing specific response tendencies relative to changes in attention to dimensions. To test this idea, they gave kindergarten and second grade children a series of two nonreversal shifts so arranged that the second shift returned to the same dimension that had been relevant for original learning. For half of the children at each grade level the second shift was a repeat of original learning; for the remainder the correct and incorrect attributes within the original dimension were reversed. This manipulation had no effect on the final performance of the second graders. But the kindergarteners learned the final problem more slowly if the relevant attributes were reversed. This is exactly the pattern of results expected on the basis of the assumption that the younger children changed their specific response tendencies more slowly than their observing responses to dimensions.

Medin (1973) and Zeaman and House (1974) have demonstrated that two-process attentional models can give a reasonable account of most of the age-related changes in children's performance on the changed and unchanged subproblems of nonreversal shifts. Consistent with the Dickerson *et al.* (1972) analysis, quantitative estimates of the response learning and dimension switching parameters of the attention model put the burden of developmental differences in the response-learning process.

Of course, there is not universal agreement that Zeaman and House are correct in their interpretation of ontogenetic performance shifts (see also Cole, 1973; Tighe, 1973; Zeaman & House, 1974). What is clear is that sharpening the question of ontogenetic changes goes hand in hand with the development of better theory.

A particularly good example of progress relating to a theory grounded in a comparative framework is to be found in the work of Sutherland and Mackintosh (1971), who have worked within a theoretical framework similar in many respects to the Zeaman–House attentional model. Sutherland and Mackintosh suggest that, in the course of mastering a discrimination, subjects learn to "switch in" relevant analyzers, which corresponds to attending to relevant dimensions. The main practical difference is that the Sutherland and Mackintosh theory has been applied mainly to the learning

and transfer of discriminations in nonhuman species. Like Zeaman and House, Sutherland and Mackintosh can interpret age and species differences in transfer experiments in terms of different rates of change in universal processes. However, they have also extended their theory to analogous but different problems in a way that extends its scope and challenges other theories to follow. We shall take a few paragraphs to illustrate their instructive use of interspecies difference in relation to their theory.

Although continued training on a discrimination should strengthen response tendencies, occasionally investigators have found that extra practice or "overtraining" facilitates learning to reverse a discrimination. This overtraining–reversal effect has been observed in rats under some circumstances (for example, difficult discriminations, large rewards), but overtraining always seems to retard reversal learning in birds. Sutherland and Mackintosh have speculated that this species difference does not arise from failures to equate all the relevant factors such as deprivation, incentives, and response requirements. Rather they propose that during overtraining, birds, unlike rats, show a tendency for response learning to be strengthened more rapidly than attention to dimensions.

Support for this explanation has come about through systematic variations in the basic paradigm that maintain close contact with the theory. For example, if original learning is followed by extinction and then a reversal, birds should show strong preferences (compared with rats) for the originally correct stimulus during the first stages of reversal learning and these preferences should be stronger after overtraining. This conclusion follows because, during extinction, birds would start to attend to other dimensions before their preference for the specific correct cue has been eliminated. An empirical confirmation of this prediction by Mackintosh (1965), combined with converging observations in the same setting, increases the plausibility of the idea that analyzers are switched more rapidly in birds than in rats.

With this success in hand, it is then possible to extend the line of explanation to other, more different, paradigms. One such paradigm involves consecutive discrimination reversals. Rats were expected to show more improvement on this task than birds, since performance should depend upon the ability to maintain attention to the relevant dimension. Slow changes in attention to dimensions (characteristic of rats on this task) should facilitate maintenance of attention to specific dimensions during a series of nonrewarded trials. Mackintosh (1969) observed that rats do perform better than doves on successive reversals. On the other hand, however, doves performed better than rats when Mackintosh presented a *nonreversal* shift at the end of a series of reversals. Again, this is consistent with the idea that rats attend more consistently to the relevant dimension than birds during the series of discrimination reversals. Finally, under a procedure

in which consistent attention to a specific dimension would *not* facilitate performance (a series of nonreversal shifts), no difference in the performance of rats and birds was observed.

In yet another paradigm, involving probabilistically reinforced discrimination, rats reach a higher asymptotic level of performance than birds. Mackintosh (1970) has shown that this difference is not the result of differences in momentary preference for the less often reinforced cue, but rather is derived from differences in the likelihood of momentary attention to another (irrelevant) dimension.

All of the above comparisons are consistent with the idea that the same two-stage attention theory can account for the discriminative performance of rats and birds, and that the differences observed are differences in the rates of changing attention to dimensions and specific response strengths. If the results from the overtraining studies had turned out to be theoretically inconsistent with those from serial reversal learning, then one might begin to question the idea that a common mechanism (attention to dimensions) was responsible for performance in the two situations. Sutherland and Mackintosh exploited the fact of species differences to test their hypotheses that a mechanism common to birds and rats was, in fact, operating in overtraining, serial reversal, and probabilistic discrimination-learning situations. The favorable evidence obtained increases the scope and power of their two-stage attention theory. At the same time we obtain a much clearer view of the nature of rat–bird performance differences.

We believe that there is a great deal to be learned from the way in which these studies were conducted. Great care was taken to make the experimental conditions for the interspecies comparison equal in so far as possible for factors considered important in performance. But this was only the beginning. Observed species differences were then placed into a theoretical framework which specified a potential basis for the species differences. After the plausibility of this interpretation had been checked to see if it was internally consistent with other observations from the original task, implications were then deduced and assessed in a range of different paradigms. In the course of this process, conditions were generated that alternatively produced and eliminated species differences in a pattern consistent with the theory. Preservation of expected relationships provided strong support that the theory had captured the sources of the species differences observed.

This procedure illustrates how the search for generality of a theory leads to the study of the behavior of different groups in different situations. Observed group differences are used to test the hypothesis that a common factor (in this case attention to dimensions) underlies performance in a number of different situations. Adequate accounts of the variability increase confidence in the scope and power of the theory and provide an

underlying framework for understanding or interpreting any performance differences observed.

For the future we hope that there will be still greater contact between the various subsets of data falling under the rubric of selective learning. For example, we would like to see how well the Sutherland and Mackintosh theory can be applied to and account for behavior in some of the mediational paradigms used by the Kendlers.

IV. CONCEPT LEARNING AND MEMORY

A currently active area of research in cognitive psychology which has close historical ties with the study of selective learning is concept learning (or concept identification). In a concept identification procedure, subjects typically are presented with several stimuli that differ along a number of dimensions (for example, shape, color, size). The task is to classify these stimuli according to some rule determined by the experimenter.

Most current theories of adult concept learning conceive of the subject as entertaining hypotheses about the solution. In this context, a hypothesis is defined as a response pattern consistent with a specifiable rule for partitioning the stimuli. Learning is represented by a process leading to the exclusive use of the correct hypothesis. Such learning logically involves developing the appropriate stimulus descriptions, generating a set of hypotheses, and applying these hypotheses to the stimulus descriptions in an appropriate way. As Falmagne (1974) points out in a review of this area, concept identification research has concentrated almost exclusively on the selection process operating on a set of prespecified hypotheses.

Within this narrow domain an important question concerns how selection of new hypotheses depends on information from the preceding trials when a hypothesis is disconfirmed. The possibilities include zero memory (information from none of the preceding trials being used, for example, Trabasso & Bower, 1968), local consistency (a requirement that the new, selected hypothesis be consistent with the just preceding trial, for example, Levine, 1966) or the idea that subjects use a focusing strategy such that all preceding events may influence hypothesis selection (e.g., Falmagne, 1972). One reasonable way to attack the problem of how information from prior events influences hypothesis selection is to work with populations who can be expected to differ with respect to their mnemonic representations of past stimuli. Before attempting such comparisons, we need a good account of hypothesis testing mechanisms in concept identification paradigms.

The starting point for such an analysis is a technique for identifying which hypothesis is controlling behavior. One technique, the exploitation of which has led to a deeper understanding of hypothesis selection, was

developed by Levine (1966) for use with college students. Working with geometric stimuli varying in such dimensions as size, color, and physical location, Levine found that by careful sequencing of the order of stimulus presentations, he was able to insert a set of probe trials between regular trials in a manner that would allow him to identify the hypothesis the subject was using.

Levine's procedure rests on the fact that if a set of eight stimulus cards vary in four dimensions, it is possible to so arrange the sequence in which the cards are presented that a consistent response pattern associated with each of the possible dimension–attribute combinations contained on the cards can be identified. Levine denotes consistent choice of a specific dimension–attribute combination as a hypothesis.

For example, in the studies to be discussed below, children are presented two cards and asked to choose "the correct one." The stimuli displayed on the cards are block letters of the alphabet that vary in their color, size, which letters are presented, and in some cases whether a particular letter is always presented on the left- or the right-hand side of the stimulus display, or has a heavy line above or below it.

The children are always explicitly told about the relevant aspects of the problem—what attributes and dimensions might possibly be correct, and are then given some practice problems. After this warmup, the main "hypothesis testing" part of the experiment begins. The basic idea of this procedure is to take advantage of the special sequencing of stimuli to give the subject information about the correctness of the choice on one trial and then "track" his hypotheses through the three probe trials, which are diagnostic because of the way the stimuli are arranged. On the probe trials, the subject makes choices as usual but is not informed as to whether they are correct or incorrect. Consequently, it is assumed that no learning occurs and that these trials simply allow an opportunity to reveal any hypothesis which he has previously selected. It is then possible to ask if a subject gives up a hypothesis if he is told he is wrong, if he keep hypotheses if told he is correct, and in general to see how choice of a hypothesis depends upon past information.

An extensive developmental investigation of hypothesis behavior using the Levine probe techniques was carried out by Gholson, Levine, and Phillips (1972). They were addressing themselves to researchers who had used this technique (Eimas, 1969; Ingalls & Dickerson, 1969), and who had concluded that hypothesis testing was not present in very young children, but emerged as a stable behavior sometime between the second and fourth grades.

In their study, Gholson and his collaborators worked with subjects ranging from seven years of age through college age. Each subject was presented a series of four-dimensional concept problems preceded by training series in which the procedure of inserting blank probe trials was introduced.

Gholson and his associates were quite successful in getting all of their subjects to respond systematically during the blank trials, but they observed a marked difference in the kind of hypothesis that subjects of different ages used. The youngest children (kindergarteners) responded almost entirely in a stereotyped fashion (such as always choosing a particular color), even though this solution had just been disconfirmed on the prior informative trial. Second- through sixth-grade students responded in a manner that Gholson and his co-workers identify as "strategic"; their behavior would, in principle, always lead to solution of the problem. However, these subjects rarely used the most efficient strategy of focusing. Finally, the performance of the college students was characterized almost entirely by a "focusing" strategy in which the subject chose hypotheses that were consistent with all of his past information.

Gholson *et al.* (1972) concluded, on the basis of their findings, that there is a qualitative change in the kind of conceptual behavior as children grow older. Their proposals for the grouping of hypotheses according to their typology represents an expansion of the theory to accommodate these qualitative changes.

Several problems in further development of the theory are acknowledged by Gholson *et al.* (1972). One such problem involves the assumption in the theory that there is no forgetting over the set of probe trials used to assess the hypothesis being used. As they note, this assumption is probably unreasonable for young children (there are data in their own as well as earlier studies to question it). One would like a direct means for assessing the extent to which changes in memory abilities between groups are responsible for the changes in the qualitative nature of the children's concept-learning performance taken as a whole, since the strategy selected may not be independent of memory factors.

A partial step toward providing the needed evidence is contained in a study by Eimas (1970), which supports the proposition that differences in memory, not differences in the ability to use hypotheses, are the source of age differences in Levine's concept learning task. Working with second graders, Eimas designed a study that provided different kinds of memory and coding aids to children in different groups. He also included a group run using the standard probe-trial procedure. In the memory aid conditions, the outcome-trial stimuli were left in the child's view during the ensuing probe trials. When children were provided with memory aids, their conceptual focusing approximated the theoretical optimum, leading Eimas (1970) to conclude that, "a major portion of the difficulty that developmentally younger children have in processing information efficiently pertains to a deficiency in memory rather than to an inability to use the logical rules of recoding and intersection [p. 331]."

This idea about the operation of memory in the concept solving process was further supported in a quantitative model of memory and concept learning published by Richard, Cauzinille, and Mathieu (1973). Using a Levine blank trials procedure with 10–12-year-olds and adults, Richard *et al.* (1973) developed a simple Markov model with three parameters corresponding to three critical processes in the concept-learning process: the probability that a subject does not try a new hypothesis when his hypothesis has been confirmed, the probability that a subject does try a new hypothesis when his hypothesis has been disconfirmed, and the probability, during the probes, that he remembers the last outcome stimulus.

Richard and his collaborators used the first two sets of probes to estimate the sampling and forgetting parameters for these trials, and then used the resulting estimates to obtain a prediction of forgetting on the third probe set.

The quantitative results are consistent with the idea that memory, rather than hypothesis testing ability, distinguishes children and adults. Both the children and adults were relatively consistent in retaining a correct hypothesis or rejecting a disconfirmed one; but substantial differences were obtained in the memory parameters. The adults were estimated to have excellent retention of previously rejected hypotheses, while the memory parameters for the children decreased rapidly and markedly.

Taken together, the results of Richard *et al.* (1973) and Eimas (1969, 1970) suggest the usefulness of comparative research for disentangling subprocesses important for concept learning. Restricting the research to adults would mean, in effect, keeping one subprocess constant at essentially a perfect level. Variations in age at first suggested variations in hypothesis testing mechanisms, but these qualitative differences themselves depended on variations in the previously constant component (memory) in the overall problem solving process. Ideally we would want a concept identification experiment that systematically varies memory load as well as age since Eimas study did not vary age.

It is conceivable that much of the variability in adult concept identification performance similarly is attributable to variations in memory and not to variations in analytical reasoning abilities. At the very least, we are aware of another critical factor in performance should we be asked to apply our understanding of concept identification processes to nonstandard subject populations.

While theoretical models of quite sophisticated nature have been developed for concept identification paradigms, there has been little empirical or theoretical work in concept-learning situations in which subjects are not provided with a set of hypotheses and the set of appropriate stimulus descriptions. Comparative research might be used effectively in exploratory

studies on these topics to provide some clues concerning the major sources of variability that one might desire to incorporate into a more comprehensive theory of concept learning.

V. CONTROL PROCESSES IN MEMORY

The modal attitude in recent developmental studies of free-recall memory is apparent in the organization of Jablonski's (1974) review of free-recall studies that have used children as subjects. Jablonski divides his paper into two sections: a review of procedures and theories derived from studies of adults and a discussion of how these procedures and theories apply to free recall in children. As Jablonski (1974) puts it: "The major theoretical issues in free recall have developmental implications [p. 524]."

One reason for the rather asymmetrical interaction between studies using adults and developmental investigations apparent in Jablonski's review is that the latter investigations have often not been conducted in any precise theoretical framework. Just the general observation that free recall performance increases with age can hardly be expected to have implications for theories of free recall. However, we shall try to indicate in this section circumstances under which a comparative investigation would be quite useful.

In the free-recall paradigm, subjects are presented with a list of words and then asked to recall them in any order. The pattern of recall provides clues about subjects' organizational activities (see Tulving & Donaldson, 1972, for an extensive review).

Some aspects of free recall relate closely to the input sequence. The probability of recalling a word is high for items at the start of a list (a primacy effect), is considerably lower for the middle portion of the list, and then is again high for the end of the list (a recency effect) producing all told a bow-shaped, serial-position curve.

One of the most influential interpretations of the serial-position effects in free recall is provided by the Atkinson and Shiffrin (1968) model for human memory. They propose that memory can be conceived in terms of three distinct memory stores: a sensory register, a short-term memory store, and a long-term store. Information is assumed to enter a short-lived sensory register from which information may be recoded in a form usable by short-term memory. Further processing in short-term may allow information to enter long-term or permanent memory.

Atkinson and Shiffrin also distinguish between the structural components of memory (such as the sensory register) and "control" processes which operate within these structural constraints. Control processes such a rehearsal may serve to maintain and transfer information from one store

to another. A computer analogy would equate structural processes with hardware and control processes with software.

The Atkinson and Shiffrin theory explains the serial position effect in free recall as follows: the primacy effect reflects the fact that the initial items in a list are rehearsed more often than other items and this extra rehearsal builds up greater strength in long-term memory. The recency effect results from an automatic "readout" from short-term memory.

In a number of studies, Atkinson and his students (Atkinson & Shiffrin, 1968; Rundus & Atkinson, 1970) found evidence consistent with the hypothesis that the principal functions of rehearsal of incoming information are to maintain it in short-term storage, to mediate short-term recall, and to promote transfer of information to long-term memory.

How might comparative research relate to the Atkinson and Shiffrin model? At the simplest level, we can ask whether children of different ages are equally likely to rehearse spontaneously when presented a set of items to remember. This is exactly the question posed in two studies by Flavell and his students (Flavell, Beach, & Chinsky, 1966; Keeney, Cannizzo, & Flavell, 1967).

Flavell discovered that spontaneous rehearsal activities in experimental situations *do* change with age: children 4–5 years of age were rarely observed to rehearse. Even many first graders (6–7-year-olds) failed to manifest any overt signs of rehearsing (such as lip movements), but those 6–7-year-olds who did rehearse remembered more than those who did not. Furthermore, when the nonrehearsers were asked to whisper the names of the objects they needed to remember, their recall was enhanced.

Thus, as a first step, we have seen that manifest rehearsal may be absent in very young children; within age groups, overt rehearsal and recall are closely associated.

If 4–5-year-old children do not rehearse, then the conditions seem ripe for some "natural" experiments concerning rehearsal and recall. Since, for example, Atkinson and Shiffrin assume that primacy in free recall results from differential rehearsal, one might expect 4–5-year-olds not to show primacy effects. Likewise, if this interpretation is supported, one can examine serial position effects obtained in related paradigms (for example, serial learning, probed serial recall) to see if a common mechanism (differential rehearsal) is in each case responsible for primacy. We can also induce different kinds of rehearsal to see what consequences they have. As we shall see, these possibilities have been explored, but a systematic picture has not yet emerged to influence significantly current theories of control process in memory.

Data on primacy and recency effects in children provide a somewhat confusing picture. In an early position-probe study by Bernbach (1967) it was reported that labeling of individual items (colors) by young children

increased both the recency and primacy sections of the recall function. The increase in primacy was extremely limited and of uncertain statistical reliability. Calfee (1967), in an extensive and thoughtful paper that included data from several studies of position-probed recall, found primacy much stronger than recency even for very young children and retardates. Hagen and Kingsley (1968) conducted a study including statistical controls suggested by Calfee and found *no* primacy but pronounced recency under conditions presumably comparable to Calfee's.

An additional manipulation was incorporated in a study by Kingsley and Hagen (1969). They compared two kinds of induced, overt rehearsal. In one case nursery schoolers were induced to label each of six nonsense figures as they were presented one at a time; in a comparison group subjects were required to label the stimuli cumulatively. One-at-a-time rehearsal did not influence primacy (vis-à-vis a control group) but increased recency. Cumulative rehearsal increased both primacy and recency. These results are potentially important theoretically because they suggest a distinction between types of rehearsal activity.

Developmental work on rehearsal has not had the impact that it might, because there simply has not been the kind of followup required to resolve discrepancies in the results of different studies and to pursue the implications of one set of findings into theoretically related areas. The problem of replication arises clearly when we consider the studies we have just been discussing. Presumed replications, upon closer inspection, turn out to contain procedural variations that could be expected, on theoretical grounds, to influence the outcomes differentially. For example, Bernbach used unfamiliar colors and asked children to point to the target color on a disk containing all four alternatives; Calfee held up target pictures using colored line drawings of animals as stimuli. Kingsley and Hagan (1969) used nonsense figures and a sample choice each trial. Hagan, Meacham, and Mesibov (1970), in a related study, used eight drawings with two black and white figures drawn on each card.

About all that can be said at this point is that neither primacy nor recency are automatic products of performance in studies of probed recall in young children, but both can be induced under some, poorly specified circumstances. Our interpretation of the different patterns of results in different experiments is hindered because of the lack of continuity between studies, and the lack of logical progression of the sort exemplified by the work in discrimination learning. But this need not happen.

The benefits of working through variations on a basic paradigm are illustrated by the work of Ann Brown, who has conducted an extensive series of studies on memory in normal and retarded children (Brown, 1975a, b), in which rehearsal was one of several control processes at issue. One line of Brown's research centers around two tasks: recognition for pictures

and a task that required subjects to keep track of the latest instance of an item seen in one of several possible categories (for example, cat was the last animal I saw, car was the last vehicle). Several studies had suggested that recognition for pictures does not require any special rehearsal or other control processes, whereas the keeping-track task requires not only rehearsal, but directed forgetting and possibly other control processes as well. In a coordinated series of studies with variations on these two tasks, Brown showed that recognition performance among mildly retarded adolescents was excellent (90% correct and better) under a rather wide variety of conditions. Performance on the keeping-track task, however, was markedly poorer in retarded than normal adolescents. She then sought to test her theory of the processes underlying performance on the keeping-track task by inducing better performance in the retarded adolescents. Her training consisted of two sessions in which the retardates were taught to rehearse overtly the current exemplar in each category. Both the pattern and the level of performance of the trained subjects changed in ways that approximated normal performance. For example, the untrained subjects showed a decrement if there were several possible instances within each category, but the trained subjects (who had been taught to rehearse—and therefore concentrate on—the last alternative) were equally proficient no matter how many exemplars there were per category. These and other aspects of the data (such as patterns of response latency) led Brown to conclude that rehearsal operates in a specifiable manner in this task and that it is a rehearsal deficit that is disrupting the retardates' performance.

This line of research was then extended in two important directions. First, the subjects were brought back several months later to see if the induced rehearsal strategies were still present; they were. Then procedures were adopted in a study with normal adolescents to see if by disrupting rehearsal, the level and pattern of retardate performance could be simulated; it could be.

Brown's work can increase our understanding of control processes in recall not only for the retarded but for normal adolescents as well because she pinpoints specific activities that underly normal recall, and then shows how the disruption of the activities has specified detrimental effects on that recall.

The firm foundation based on the keeping-track task can now be used to explore the role of rehearsal in a variety of other memory situations where rehearsal is presumably required.

It might be noted that in Brown's research, as well as the work of Flavell, Calfee, Hagen, and others discussed here, the major emphasis has been on specifying differences in the "control process" used by different subject populations, rather than emphasis on different structural properties of memory (for example, none of these researchers has attributed group differences to differences in the size of a short-term memory buffer).

In an article that already has had considerable influence on comparative studies of memory, Craik and Lockhart (1972) pursue this emphasis on control processes into the adult literature on memory. They suggest that the "multiple store plus control process" approach be supplanted by a theory of memory that relies almost entirely on a specification of the nature of the control processes. According to this approach, the idea of distinct memory stores should be supplanted by the characterization of memory as a "byproduct" of information processing activities which may or may not have remembering as an end result.

Craik and Lockhart (1972) put the matter as follows: ". . . it is suggested that the memory trace is better described in terms of depth of processing or degree of stimulus elaboration. Deeper analysis leads to a more persistent trace [p. 677]."

Restle (1974) generalizes the Craik and Lockhart position: "Encoding is the process of incorporating information into a cognitive structure, and the depth of processing corresponds to the generality and size of the cognitive structure involved [pp. 216–217]."

The depth of processing approach orients us toward a more careful look at the classes of activities that have been grouped together under the rubric of control processes in general and rehearsal in particular. In his own work, Craik has concentrated a good deal of attention on the different activities that we have been discussing here as rehearsal. He makes the distinction between rehearsal activities that *maintain* information in an accessible state and activities that *"enrich and elaborate"* the information for later retrieval. As Craik and Lockhart (1972) put it, "whether rehearsal strengthens the trace or merely postpones forgetting depends on what the subject is doing during rehearsal [p. 681]."

The majority of work on the depth of processing approach has relied on experimental manipulations that limit the information processing activities of adults and then assess the consequences for recall. Related studies with adults (e.g., Jenkins, 1973) show that incidental recall can equal or exceed intentional recall of meaningful material depending upon how the subject is asked to deal with the stimulus materials.

While this strategy has produced a number of interesting results, there may be some unrealized potential for "depth-of-processing analysis" in the use of nonstandard populations. That is, there is a further implication to the depth of processing approach which is made explicit in Restle's comments quoted above and which seems to have received relatively little attention, so far as we can tell. It is assumed in the adult work that by manipulating instructions and other aspects of the way in which tasks are presented to the subject, one can manipulate the cognitive structures involved. But when we think of work with children or retardates, there is another, natural way to manipulate the size and generality of cognitive structures because

these populations are generally assumed to have less developed, less general cognitive structures. We see considerable promise in adding comparative research to the set of tools used by cognitive theorists interested in exploring the depth of processing approach precisely because an additional source of theoretically relevant variation is added. A natural question to ask is how such processes as cumulative rehearsal and intentionality interact with subjects' ability to understand and interpret the materials as well as their readiness to use certain information processing procedures.

What seems called for are studies that look at the interaction between depth of the processing activities and the depth of subjects' understanding of the materials they are being asked to deal with. As a concrete example, suppose that we presented pictures to children for recall in a position-probe task of the kind used by Calfee and Hagen. Suppose that we selected the materials so that a younger group of children could, potentially, understand the structure of a series like cat–mouse–dish–cheese but not an alternative series such as bride–ring–groom–rice. An older group that could comprehend the structure of both series would also be studied. The younger children might be induced to rehearse each series (through naming) in a cumulative or one-at-a-time manner. Such an experiment might serve to clarify the relation between "depth" caused by different rehearsal activities from "depth" resulting from differential comprehension.

We hope it is clear that we are not arguing that comparative research is the *only* way in which to test the cognitive implications of the depth of processing approach. We are simply making the case that comparative studies offer an approach to the study of control processes which should not be overlooked.

VI. MASSED AND DISTRIBUTED PRACTICE

In the early 1950s James Deese commented that verbal learning theorists would find little new or interesting information in the literature on children's learning (Deese, 1952, p. 297). This thought was echoed about a decade later by Keppel (1964) and it seems safe to say that it represents the opinion of many research workers in the field today. It is something of a challenge, then, to see if comparative studies of verbal learning in recent years give any cause to question this negative judgment.

As an illustration, we have picked the hoary topic of the effects of massed and distributed practice on recall. Keppel (1964), operating on rather sparse data, concluded that the spaced-practice effect is probably similar in adults and children. He could cite only a single study, conducted in 1926, to support his conclusion and pointed out that Underwood and his colleagues (Underwood, 1961) had found distributed practice to have more influence on retention than on learning of paired associates.

In the past decade, research on paired-associate learning in adults has shifted emphasis from the learning and retention of whole lists to a concern with the fate of single items after a single study and test sequence. The shift in method has been accompanied by a corresponding shift from the claim that spacing effects are small and unreliable (Underwood, 1961) to the current conclusion that the advantage of spaced presentations is quite large and extremely robust (Hintzman, 1974).

The advantage of spaced presentations seems to be so general that it must be telling us something important about memory, and it is not surprising that a fair number of theories designed to explain this effect have evolved. But the very robustness of the spacing effect has impeded theoretical progress—it is hard to make the effect go away, at least with the usual subject populations and paradigms.

Suppose the effects of spaced repetitions on memory are different or perhaps even absent, in other subject populations. Have theories of the spacing effect evolved to the point where we are in a position to interpret differences in its magnitude across subject populations? We suspect that, aside from some tentative speculations, the answer to the latter question would be a grudging, "no."

Against a background of hundreds of studies on spacing of repetitions in college students, there are only a handful of experiments using any other subject population. Nevertheless, we shall try to garner an example or two of how understanding of the spacing effect might benefit from comparative research.

Although spacing of repetitions generally seems to help performance, in some paradigms (for example, paired-associate learning) the spacing effect is nonmonotonic. That is, initially as spacing increases, performance increases; but the spacing advantage reaches a maximum, and for long spacing intervals, performance tends to decrease toward the level observed for massed repetitions (e.g., Peterson, Wampler, Kirkpatrick, & Saltzman, 1963).

There is no fully satisfactory account of this nonmonotone spacing effect. One prominent interpretation of the effects of spacings of repetitions in college students uses the distinction between short-term memory (which is subject to rapid forgetting owing to decay or interference) and long-term memory (which is assumed to be resistant to forgetting). The relevant further assumption is that when an item is already represented in short-term memory, additional presentations of that item are not effective and do not promote transfer to long-term memory. Spaced repetitions are more beneficial than massed repetitions because at longer intervals between repetitions it is less likely that the representation of the first presentation will still be in short-term memory at the time of the second presentation. Models of this form (e.g., Atkinson & Shiffrin, 1968) have done a fair job of

accounting for observed spacing advantages (see Bjork, 1970, for a review).

A weak link in this formulation is that it fails to explain why the spacing effect in paired-associate learning reaches a maximum and then declines. Theories of the Atkinson–Shiffrin variety could be modified to explain the nonmonotonic spacing effect by adding the assumption that there is some probability of forgetting an item from long-term memory. According to this line of reasoning the first presentation of widely spaced items would be more likely to be lost from long-term memory than the first of two massed repetitions, since longer retention intervals would be involved in the former case. Consequently, the spacing effect would result from a trade-off between the ineffectiveness of a repetition when an item is already in short-term memory and the deleterious effect of widely spaced repetitions owing to the greater chance of forgetting from long-term memory.

Data from recent experiments by Sperber, Greenfield, and House (1973) and Sperber (1974) provide a test of this augmented short-term/long-term memory theory, and a nice illustration of the way in which a comparative experimental design can lead to new theoretical developments.

A key point in our provisional explanation of a nonmonotonic spacing effect is that items may be lost from long-term memory. Sperber *et al.* (1973) conducted a study using a verbal discrimination paradigm quite analogous to the familiar paired associate technique with retarded children. They arranged their procedures so that they could present 1, 3, 4, or 8 trials per day. When more than one trial was presented on a day, the trials were massed. The one-trial-a-day condition represents the greatest spacing interval, of course, and Trial 9 is a critical trial for testing the spacing effects since the ninth trial follows a constant 24-hr interval in each of the presentation conditions.

The spacing effects obtained by Sperber and his collaborators were nonmonotonic: it was found that 2 and 4 trials per day yielded the best performance, whereas 1 and 8 trials per day resulted in consistently poorer performance on the ninth trial. Using 24- and 72-hr retention tests, Sperber *et al.* (1973) found very little forgetting. There was so little forgetting, in fact, that resort to the notion of forgetting from long-term memory could not give an explanation of the nonmonotone spacing effect.

In a related study Sperber (1974) compared low mental age (MA) and high-MA retardates and found spaced repetitions to be better than massed after long retention intervals for high-MA subjects, but he found a uniform advantage for massed repetitions in low-MA subjects. The two groups did not differ reliably in their overall performance, nor on the second trial of problems, and so differences in learning rate cannot explain the difference in performance. Neither did the two groups differ on 24-hr retention when the original learning trials had been massed. These results add

yet another problem to the "forgetting from long-term memory" theory, which must be able to predict the absence of a spacing effect (or a massing advantage) in low-MA retardates, but a nonmonotone result from high-MA retardates.

To account for these data, Sperber *et al.* (1973) proposed that population differences in an initial coding process accounts for their results. The main idea is this: when an item is initially presented, it is unfamiliar (at least in the particular experimental context). Forming a stable long-term representation or effective code of the item is more difficult and less likely when an item is unfamiliar or uncoded than when it is familiar and coded. What is new in the Sperber *et al.* (1973) theory is the proposition that over the course of a retention interval the initial encoding or familiarity might be lost. The processes involved are imagined to function as follows: massed trials are ineffective because there is great overlap of information already in short-term memory, but at very long intervals, loss of a distinctive code offsets the absence of redundancy and forgetting is increased.

These ideas were formalized by Sperber *et al.* (1973) as a mathematical model which they applied to their data. The model represents a fairly straightforward modification of a set of Markov models known as the general all-or-none forgetting theory, which has been analyzed and applied to adult data (Bjork, 1970). The modified model accounted for over 91% of the variance in the data.

In short, the idea that loss of coding occurs was able to account quantitatively as well as qualitatively for the nonmonotonic spacing effect. The difference between low-MA and high-MA retardates noted by Sperber (1974) could be handled by assuming differences in the probability that the distinctive code would be lost: low-MA retardates were much more likely to lose "familiarity" than high-MA retardates. Other parameter estimates indicated that a repetition was not effective when an item was already in short-term memory (an assumption often supported in Bjork's research using college subjects).

Overall then, Sperber and his associates were able not only to make sense of their data, but also to shed considerable light on processes which may be more generally operating in memory of adults as well as children. The results of Sperber and his colleagues (Sperber, 1974; Sperber *et al.*, 1973) suggest that the initial coding of a stimulus plays a major role in memory. The possibility arises that retardates have not developed the cognitive structures to allow them to form an effective encoding of an event. This line of thinking might provide a link between the concept of an initial coding and the "depth-of-processing" framework outlined by Craik and Lockhart. However, before engaging too much in speculation, it is worth considering how the results we have just been discussing might be made even more convincing, especially from a comparative point of view.

First, if the nonmonotonic spacing effect results from loss of initial coding, then an important experiment would be to try to manipulate the likelihood of code loss and to show corresponding changes in the form of the spacing effect. By varying properties of the learning materials to produce variability in the probability of an initial code being formed and the probability of its being retained we might be able to produce massing advantages or spacing advantages in both high- and low-MA retardates, depending on the particular combination of experimental conditions. The procedures for establishing initial encoding as the source of the experimental effects give a better indication of its importance in a general theory of memory and go a long way toward elucidating the comparison of high and low MA performance.

Clearly, more work is needed to determine just what this initial encoding or familiarity involves and how it might be lost. Until we looked closely at retardate learning, we had no reason to expect that loss of initial encoding might occur or that it might be responsible for nonmonotone spacing effects. With a restricted subject population in a restricted set of paradigms, there may be "hidden" source of variability and a good fit of a model to such data may belie a lack of understanding of the cognitive factors operating ubiquitously.

VII. FREE RECALL AMONG A NONLITERATE PEOPLE

Early in this chapter we made the claim that comparative cognitive research encounters similar problems and promises similar potential benefits to cognitive theory across a very broad spectrum of possible population comparisons. Thus far, we have restricted our examples to ontogenetic, phylogenetic, and psychometric IQ comparisons. In this section we would like to consider some of the problems and benefits of cognitively oriented, cross-cultural work.

The particular series of studies we will describe was begun in rural Liberia in connection with a project designed to improve the education of rural Liberian children entering American-style schools in the hinterland (Gay & Cole, 1967).[1] Our interest soon spread to the related question of how schooling experience influences cognitive development. This orientation is important to consider because it influenced the choice of populations studied (educated and noneducated).

We do not attempt to describe this entire project (see Cole, Gay, Glick, & Sharp, 1971). Rather, we give a schematic account of the subset of that project which dealt with differences between educated and noneducated

[1] In order to simplify the text, this work will be described in the first person plural. It should be understood that "we" and "our" refers to Cole and those of his colleagues specifically cited.

Liberian (Kpelle) people in their performance on multitrial free recall tasks.

One phenomenon that caught our attention is illustrated by the free-recall performance graphed in Figures 1 and 2. Figure 1 shows the number of correctly recalled items as a function of trials for two groups of Kpelle teenaged subjects presented a clusterable list of 20 words representing four taxonomic categories (clothing, food, tools, utensils). The group labeled "High School" was made up of students attending a rural high school; the "Village" group was made up of young adults from a nearby town where none of the subjects had attended school. In this and a companion study (see Cole *et al.*, 1971, p. 129) it is clear that the High-School subjects begin by recalling slightly more items than their noneducated counterparts and recall performance of the former group improved over trials. The performance of the Village subjects improved slightly for a few trials, but little thereafter. By the end of training there was a 4–5-item difference between the groups in the number of items recalled.

An even more dramatic difference between the groups appears in Figure 2, which presents data on the semantic clustering of free recall performance paralleling the recall performance in Figure 1. (*Clustering* refers to the grouping of items in recall that were not presented contiguously at the time the list was presented. *Semantic clustering* as the term is used here refers to grouping according to taxonomic category during recall.) There

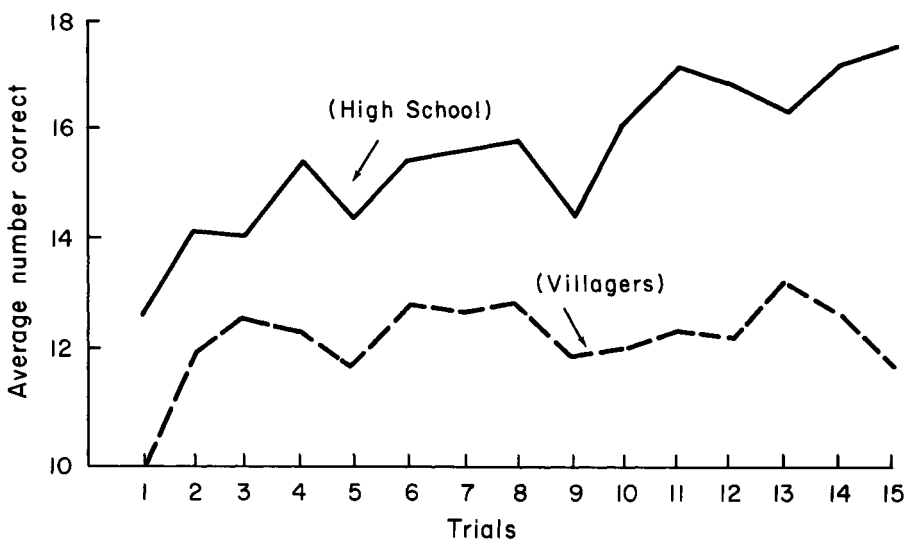

FIGURE 1 Average number of items recalled per trial for high-school and village subjects.

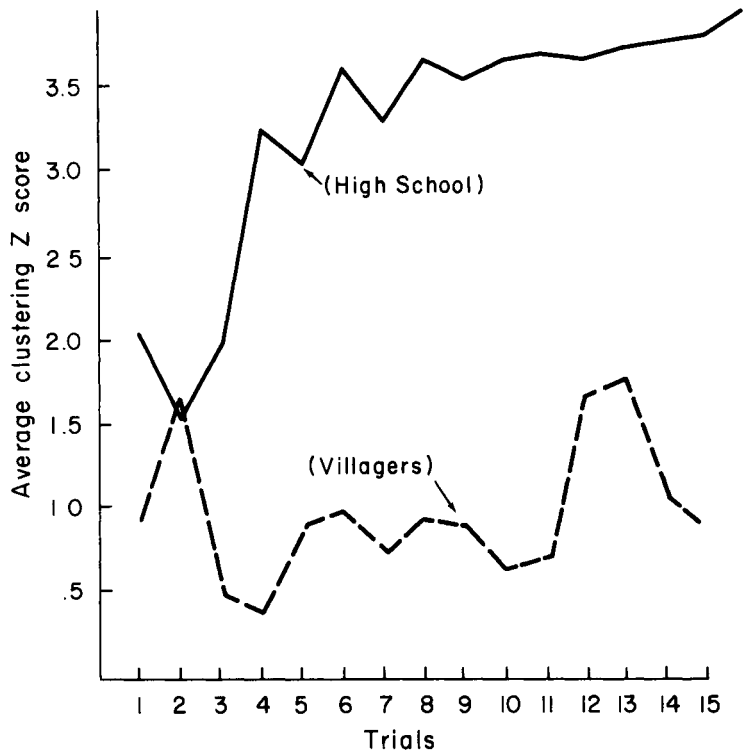

FIGURE 2 Average semantic clustering scores per trial for high-school and village subjects whose recall data are plotted in Figure 1.

is virtually no improvement in clustering for the village subjects, but rapid improvement for the High Schoolers.

Working further with these data we ruled out the possibility of trial to trial idiosyncratic consistencies in the ordering of subjects' output that Tulving (1962) has termed "subjective organization." We also tested the hypothesis, first put forward by Bartlett (1932), that nonliterate people are likely to try to rote memorize. The correlation between the order in which items were presented and their order of recall approximated zero in the villagers' recall protocols.

Why do the Villagers fail to learn the list? In fact, why is there virtually no improvement in their recall after Trial 2 or 3? Why is there a lack of the clustering in free recall, which is certainly present in the performance of the educated subjects and a virtually ubiquitous finding for normal American adolescents (Jablonski, 1974)?

One of the first factors that we want to consider is the recall materials themselves. As many investigators have emphasized, factors such as item

familiarity, association between items, and association between items and category labels all influence free-recall performance (Hall, 1971).

Operating in a nonliterate culture with no background of word-frequency or category norms to guide us, the problem of understanding the familiarity and lexical structure of the to-be-recalled materials represents a major problem. Hence, one of the first tasks in pursuing the problem of interpreting our Villagers' performance was to gain some understanding of the stimulus materials.

A variety of techniques, some borrowed from linguistic anthropology, were used to assure ourselves that we were working with familiar items, taxonomically organized according to the categories that we had assumed to constitute our categorized list (see Cole *et al.,* 1971, Chapter 3). The upshot of this work was that we were able to demonstrate to our satisfaction that the 20 clusterable items were indeed familiar, highly interassociated, and likely associates of the category names. Hence, we could rule out the possibility that differences between the groups arose from differences in subjects' knowledge of the items and categories used.

However, it might be that although our subjects reproduce the lexical structure built into our recall list when engaging in nonmnemonic tasks (free associating, naming category instances or superordinates), they would use a different organizing principle when asked to recall these same items. For example, Denny and Ziobrowski (1972) demonstrated that young children from the United States are likely to cluster items thematically rather than taxonomically, although they can (and do) name appropriate category instances if requested to do so.

An attempt to test the hypothesis was made by constructing a list of four categories of common items that were related thematically (for example, "things for catching meat" such as gun, net, hook, trap, and bow). This list produced results exactly analogous to those in Figures 1 and 2.

We then instituted a set of studies designed to explore hypotheses about other significant factors that might be controlling the Villagers' performance, some arising from anthropological sources, and some from the psychological literature.

One persistent report on the sparse literature on cross-cultural comparisons of intellectual processes is that tribal Africans think "concretely" (Cryns, 1962). While this notion of concreteness has never been spelled out, we took it seriously enough to believe that if we presented villagers with the actual objects to be remembered, instead of naming them, we would see a dramatic improvement in performance. Hence, one of our early studies was a comparison of recall for oral and object stimuli. One group of subjects was read a list of nouns at approximately a 2-sec rate; the other was shown the objects named at an equivalent rate. At the time this work was initially planned and executed, there were no such comparisons

extant in the literature on free recall in the United States, and a good deal of the work in the United States had required literate subjects to write out their recalled items. Written recall was clearly inappropriate for our nonliterate subjects, so we quickly found ourselves with an additional experiment to do (an experiment in the United States comparing recall for objects and orally presented words using oral and written recall modes) and a new cross-cultural hypothesis: the difference between object and oral presentation should be relatively greater among the Villagers than among Kpelle High Schoolers or groups from the United States at a comparable age. This enterprise produced some data of interest to those studying free recall among children from the United States (such as the fact that written recall produced more clustering and better performance than oral recall under a variety of conditions). However, we found very little difference between the presentation of objects and orally presented stimuli among Kpelle villagers, even less than for American subjects ranging from first through sixth grade. The major features of the Villagers' performance, as illustrated in Figures 1 and 2, remained—low levels of recall, little improvement over trials, and virtually no clustering in recall.

We next looked into the matter of motivation. Perhaps the Villagers were simply indifferent to their performance. After all, remembering things for a stranger from the United States was not something they had any particular reason to value. To be sure, subjects were paid a nominal fee for participating in the study, but we had no evidence that they were really trying to remember all the items.

In order to manipuate the incentive for good performance, we arranged a study in which half of the subjects were paid a penny for each word correctly recalled. Since there were 5 trials and 20 words, a total of 1 dollar could be earned in this way. This contingency was explained to a new group of subjects who were otherwise treated exactly as the Villagers whose performance is graphed in Figures 1 and 2. (Lest it be thought that $1.00 did not represent a sufficient incentive, the average daily wage rate in Liberia at the time of this study was between $.50 and $.75 per day of hard labor).

The results of this study were very clear-cut. The monetary incentive affected neither recall accuracy nor clustering. However, the paid subjects *said* many more words at the time of recall; their protocols were replete with repetitions of recalled items, as well as many intrusions of items that had not been named.

After following up several other unproductive hypotheses, we began to make some progress when we concentrated on the issue of making the categorical structure of the materials clear to the subject.

Our initial attack on this issue was to depart from the free-recall procedure in an attempt to provide a clear rule, external to the to-be-recalled

items themselves, that could serve as a basis for organizing the materials. After some pilot work, we hit upon the idea of conducting a somewhat unorthodox paired-associate experiment in which the stimuli were four wooden chairs, lined up in a row in front of the subject. The response items were the 20 clusterable objects used in the High-School–Villager comparisons. Two experimental conditions were contrasted in a study with a new group of villagers. In one condition, the assignment of objects to chairs was based upon their category membership (all clothing to Chair 1, all tools to Chair 2, etc.). In the other condition, items were assigned to chairs such that no chair contained more than two items from the same category. In both conditions, there was a different randomized order of presentation of chair–object pairs on each trial. Under these conditions, there was a clear superiority of the groups that had the items assigned to chairs on the basis of common category membership.

These results encouraged us to return to the free recall paradigm with the aim of making the categorical structure of the clusterable list as explicit as possible. A small study that added category names into the to-be-recalled list failed to have any effect. However, a pair of studies, associating to-be-recalled items with chairs (even though subjects were required only to recall the objects, not the object–chair connection), produced a marked improvemment in clustering and recall. In one study, designed to parallel the paired associates experiment, objects were held up over chairs as "reminders" in either a random fashion or in a way such that items from a specific category were always held up over a particular chair. When the items were held up at random, learning was poor and there was no categorical clustering (there was even a hint of clustering by chair instead of by category). But when items were held up over chairs in accordance with their category membership, learning curves like those obtained for the High-School subjects in Figures 1 and 2 were obtained.

Unhappily, the effects of "cuing by chairs" in free recall turned out to be considerably more complex than we had anticipated. In a replication of the study we found good recall even under the random condition, or a condition with a single chair. While it is interesting to speculate about what was producing the good recall when different cuing-by-chairs procedures were used, the main outcome of this kind of study was to convince us that villagers were capable of good recall. Our problem was that we were still incapable of specifying the conditions under which the good recall could be obtained or what processes were operating when recall did not improve over trials.

While attempting to disentangle possible features of the cuing-by-chairs procedure, we also undertook to see if we could obtain similar improvement of results by purely verbal means. We had just become familiar with the work of Tulving and his associates (Tulving & Osler, 1968; Tulving

& Pearlstone, 1966) suggesting that items might be remembered in the sense that they are stored in memory and yet be inaccessible to retrieval at the time of recall. Working with 10–14-year-old village children who had attended little school, we conducted a study modeled after those of Tulving and his associates.

The basic idea of this experiment was to cue subjects, *explicitly,* about the categorical structure of the to-be-remembered list at either the time items were presented (input) or at the time they were to be recalled (output). The four basic experimental conditions represented the factorial combination of cuing at input and at output. For example, if the subject was cued at the time of input, he was told that the list consisted of foods, tools, clothing, and utensils just prior to presentation of the list. If cuing occurred at the time of output, the subject was reminded of the four categories as soon as the list had been read.

The result of the main part of the experiment was that verbal cuing at input, output, or a combination of input and output made no difference whatsoever in the overall performance of these village youngsters; once again the patterns for performance in Figures 1 and 2 repeated themselves.

But a fifth condition produced a really dramatic change in performance beginning with Trial 1. In this condition, the subjects were cued at both input and output. In addition, their recall was constrained for the first four trials. The constraint was introduced just following the cuing at output time. Instead of simply asking the subject to recall all he could, the experimenter systematically elicited recall one category at a time. For example, a subject might be asked, "First, tell me all the clothes you remember." Then he would be asked, "Now, tell me all the tools you remember."

Under these constrained recall conditions, recall improved on Trial 1 by an average of more than five items and remained high for the four constrained trials. On the fifth trial the constraints at recall time were removed, so that the procedure was identical to that for the group which had been cued at input and output. Recall remained high, and, most interesting, there was highly significant clustering.

The dramatic improvement on the very first trial of recall strongly implicates factors working at the time of retrieval in the previously observed performance of the villagers. Their difficulty arose not so much from a failure to recognize that the items in the list belonged to categories, as from a failure to use this information in the service of recall. These results suggested several lines of investigation, only a few of which we have had the opportunity to follow up.

Our first question was whether we could discover conditions under which villagers would spontaneously use the potential structure of the stimulus input to organize their recall. One promising procedure was to embed the to-be-recalled materials in pseudofolk stories. Story telling is a popular

recreational and educational pastime among Kpelle villagers, and we hoped that by introducing materials in this context, the structure of the story would be used to structure recall.

One of the stories used was no more than a pretext for presenting our clusterable list: a chief offers his daughter's hand in marriage to a visitor if the man can remember all the objects in the chief's house; the list is read, and the subject is asked to try to recall all the items in the chief's house. This procedure produced the same lack of cumulative improvement we had observed repeatedly among the villagers with repeated trials of oral free recall. A second story structured the input by recourse to a story in which four men come to town to ask the chief's permission to marry his daughter, and each man brings five gifts (corresponding to the five items in each category). The subject was asked to name the gifts brought by each man and to argue for the appropriate bridegroom. Under these conditions recall was highly clustered, although some subjects recalled only the items of the man they deemed the appropriate groom! Two other stories strung the items out in a narrative. These conditions produced little clustering, but a strong correlation between input and output order—just as we would expect if the subject were using the story structure to mediate his recall. It thus appears that a strong narrative framework induces organization of recall. We would like to know of other conditions that would promote the same outcome.

It is one of the drawbacks of cross-cultural research that long-term continuity in the field is rare. Follow-up of this work among the Kpelle has been sporadic since these studies were undertaken in the late 1960s. However, we have been fortunate in being permitted to pursue some of the many questions that this line of research has raised in shorter term studies, and in having them replicated and extended by other psychologists.

Sylvia Scribner (1974) contributed a valuable additional piece of information in two studies that pursued the problem of the relation between subjects' own organization of to-be-recalled materials and subsequent recall. In our original series, we had included a study in which subjects were required to categorize the set of objects and then recall them, but we had arranged matters such that categorizing and recall were interspersed: categorize–recall, categorize–recall, etc. This procedure produced no noticible deviations from our base-line performance. Scribner, following Mandler and Stephens (1967), required groups of villagers and high-school students to sort a group of objects to a criterion of two identical sorts in a row. Under these conditions, subjects' recall mirrored their sorting (although villagers were less likely to sort by taxonomic category than high schoolers). These results suggest that our failures to find spontaneous retrieval by category in our earlier studies were conditioned in part by a failure to induce strong categorical organization within the remembering situation

prior to asking for recall. This line of interpretation is consistent with the results of the paired associates and constrained recall studies, but uses procedures that make closer contact with the American literature.

Although clearly incomplete, the series of studies illustrates a number of the points that we have tried to make about comparative research in the earlier parts of this paper.

The observation of a wide difference in group performance was used as the occasion to identify the subprocesses that produced the better performance of the high-school group. No simple "method of equation" and no single study proved adequate in this search. Rather, using theoretical ideas from psychology and anthropology, we were forced to undertake and extensive (and still incomplete) series of studies that systematically varied aspects of the stimulus materials and experimental conditions that we guessed might be of primary importance.

Were this search to begin today, we would not have been led into quite so many blind alleys (only a few of which have been described here). Many seminal studies, including those of Flavell, Brown, and others described earlier, would have led us more quickly to the general line of attack that eventually proved fruitful. By the same token, the fact that we have arrived at a set of explanatory concepts similar to those employed by Flavell, Brown, and others by a very different route lends extra weight to current memory theories that emphasize the role of subject-controlled mnemonic strategies in free recall performance.

Some of our observations also have led us to question theoretical formulations derived solely from laboratory studies of memory. For example, we had initially expected verbal cuing at the time we presented items to enhance recall and clustering in line with Mandler's (1966) statement that organization in recall would occur so long as the subject was permitted to discover the potential structure in the list. Our results led us to conclude that in addition to the discovering or recognizing of potential structure, some additional processes were necessary to induce subjects to use the structure in their attempts to remember.

These results may also serve as a starting point for those interested more in questions of the role of experience in the development of memory skills than in theories of memory per se. Both in our work in Liberia and in subsequent replications and extensions in rural Mexico (Sharp & Cole, 1974; Wagner, 1974) there is the strong suggestion that it is the level of educational attainment, not age alone that is important in "age" related differences in memory performance in the United States. While currently available data are far from conclusive, we are now in a much better position to answer such questions as "how does educational experience influence cognitive performance" (at least memory) than we were before these studies were undertaken.

VIII. GENERAL DISCUSSION: THE THEORETICAL AND APPLIED BENEFITS OF COMPARATIVE RESEARCH

Rather than extend this highly selective review of research with nonstandard populations, we would like to raise again the questions with which we began, and to reexamine them in the light of the examples we have been discussing.

To begin with, we can get a clear perspective on the lack of attention paid to research on nonstandard populations by many cognitive theoriests. It is not simply a matter of different subject populations, but different scientific enterprises. The modal paradigm in cognitive research with adults is to hold the subject population constant and to vary the task in a variety of ways designed to test the implications of a verbally formulated, or perhaps mathematically formulated model. Frank Restle, whose work has been very much in this tradition, characterizes the enterprise and its problems as follows:

> Experimental research results in an understanding and a theory of the task the subject is working on . . . researchers pick a task and then investigate the requirements of the task itself. . . . This means that it is important to choose an important task. I feel that we have dug ourselves into deep pits sometimes by studying a task just because it is there and has been useful at some time in the past to answer some theoretical question [Restle, 1974, pp. 223–224].

Concerning the area of verbal learning, Restle suggests that memory studies involving semantic integration of sentences are more important than list learning tasks, among other reasons, because people have to do semantic integration all the time.

One need not agree with Restle's notion of an important task to realize the significance of his assertion that cognitive psychologists study tasks, not people. The strategy followed by comparative research is in some respects the mirror image of that used in the research referred to by Restle and can be a useful complement to it: a single task (free recall, discrimination transfer, concept identification) is applied to a variety of subject populations. As a rule, differences in performance are attributed to differences in the populations with respect to the process said to underlie performance in the task: normal teenagers have developed rehearsal strategies, retardates have not; seven-year-olds have developed mediation, two-year-olds have not.

The theme of our reveiw is that a sharp "division of labor" may be counterproductive to the extent that it circumvents interaction between cognitive theory and comparative research. In order to understand the full

range of consequences of task structure *and* in order to understand the full range of differences among subject populations, we suggest that it is advantageous to develop strategies that incorporate variations in both task and subject variables within a single experimental program. The examples of research topics we have reviewed here only partially fulfill our notion of the effort needed. Nevertheless, they illustrate benefits to be gained from obtaining both task and subject variability, even when the variability is not dissected in a particularly systematic fashion.

IX. CAN WE APPLY COGNITIVE RESEARCH?

Thus far, we have concentrated our attention almost entirely on the question of whether research with nonstandard populations can be relevant to cognitive theory. This orientation has led us virtually to ignore a great deal of research with nonstandard populations conducted for quite different purposes. Here we discuss two aspects of the question of the relevance of cognitive theory to other areas of psychology, and ultimately to society.

First, we would like to return to the distinction between historical and systematic explanation that we introduced early in this presentation. We are not opposed in principle to exploring the antecedents of performance on cognitive tasks since questions about such antecedents are often considered relevant beyond the laboratory. In fact, we have been very concerned with understanding the role of educational experience as an antecedent of several of the tasks we have described in this paper (see Cole & Scribner, 1974). However, we have been very dissatisfied with the range of inferences that current cognitive theory allows us to make from particular patterns of test performance—hence our emphasis on pursuing antecedent explanations using tasks for which we have as highly developed a theory as possible.

But even in cases where we have a strong and accepted theory of constituent behaviors, we are still faced with Restle's question of how to select an important task. In our work on the effects of education, we have selected tasks using a dual criterion; a belief that we have systematic understanding of the behavior the task is sampling, and a belief that people who do well in school are likely to excell at the task.

A systematic version of this kind of relevance is to be found in the work of Ann Brown (1975a, b), who proposes that her research strategy could be used as a guide to curriculum. Beginning with the generalization (derived from her own and other studies) that retarded children experience greatest difficulty on tasks that require strategic planning, the educator would first analyze the material he wants to teach into its theoretically based strategic components. Next, he would test the retarded child to see

if he spontaneously engages in the necessary strategic behavior. In the event of a negative result, the educator would then undertake a training program to *induce* the behavior for the setting in question. Finally, means should be developed to generalize the behavior to similar settings and to insure its maintenance over time.

Still, these approaches to relevance stick pretty closely to school, the very context from which most of our tasks were derived in the first place. Moving beyond the range of tasks suggested by educational problems is unlikely to occur according to any single criterion of relevance. There are a great many tasks that people engage in often (to return to Restle's relevance criterion). We see no way of providing a content-free formula relevant to all nonlaboratory concerns. However, we do believe that the orientation toward cognitive research that we have been discussing will allow movement toward a situation in which the sources of variability that we isolate in the laboratory are the sources of variability that matter in real-world situations.

ACKNOWLEDGMENTS

Preparation of this paper was supported by a Grant from the Carnegie Corporation and Grant MH23878 from the National Institute of Mental Health.

REFERENCES

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.

Bartlett, F. C. *Remembering.* London: Cambridge University Press, 1932.

Bernbach, H. A. The effects of labels on short-term memory for colors with nursery school children. *Psychonomic Science,* 1967, **7,** 149–150.

Bitterman, M. E. Phyletic differences in learning. *American Psychologist,* 1965, **20,** 396–410.

Bjork, R. A. Repetition and rehearsal mechanisms in models of short-term memory. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970.

Brown, A. L. The role of strategic behavior in retardate memory. In N. R. Ellis (Ed.), *International review of research in mental retardation.* Vol. 7. New York: Academic Press, 1975. (a).

Brown, A. L. The development of memory: Knowing, knowing you don't know and know how to know. In H. Reese (Ed.), *Advances in child development and behavior.* Vol. 10. New York: Academic Press, 1975. (b).

Bryant, P. *Perception and understanding in children.* New York: Basic Books, 1974.

Calfee, R. C. Short-term recognition memory in children. *Child Development,* 1967, **41,** 145–161.

Campbell, D. T. The mutual methodological relevance of anthropology and psychology. In F. L. K. Hsu (Ed.), *Psychological anthropology.* Homewood, Illinois: Dorsey Press, 1961.

Cole, M. A developmental study of factors influencing discrimination transfer. *Journal of Experimental Child Psychology*, 1973, **16**, 126–147.

Cole, M., Gay, J., Glick, J., & Sharp, D. W. *The cultural context of learning and thinking*. New York: Basic Books, 1971.

Cole, M., & Scribner, S. *Culture and thought*. New York: Wiley, 1974.

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.

Cryns, A. G. J. African intelligence: A critical survey of cross-cultural intelligence research in Africa south of the Sahara. *Journal of Social Psychology*, 1962, **2**, 283–301.

Deese, J. *The psychology of learning*. New York: McGraw-Hill, 1952.

Denny, N. W., & Ziobrowski, M. Developmental changes in clustering criteria. *Journal of Experimental Child Psychology*, 1972, **13**, 275–283.

Dickerson, D. J., Novik, N., & Gould, S. Acquisition and extinction rates as determinants of age changes in discrimination shift behavior. *Journal of Experimental Psychology*, 1972, **95**, 116–122.

Eimas, P. D. A developmental study of hypothesis behavior and focusing. *Journal of Experimental Child Psychology*, 1969, **8**, 160–172.

Eimas, P. D. Effects of memory aids on hypothesis behavior and focusing in young children and adults. *Journal of Experimental Child Psychology*, 1970, **10**, 319–336.

Estes, W. K. *Learning theory and mental development*. New York: Academic Press, 1970.

Falmagne, R. J. Memory process in concept identification. *Journal of Experimental Psychology*, 1972, **92**, 33–42.

Falmagne, R. J. Mathematical psychology and cognitive phenomena: Comments on preceding chapters. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology*. Vol. 1. San Francisco: Freeman, 1974.

Flavell, J. H., Beach, D. R., & Chinsky, J. M. Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development*, 1966, **37**, 283–299.

Gay, J., & Cole, M. *The new mathematics and an old culture*. New York: Holt, Rinehart & Winston, 1967.

Gholson, B., Levine, M., & Phillips, S. Hypotheses, strategies and stereotypes in discrimination learning. *Journal of Experimental Child Psychology*, 1972, **13**, 423–446.

Hagen, J. W., & Kingsley, P. R. Labelling effects in short-term memory. *Child Development*, 1968, **39**, 113–121.

Hagen, J. W., Meacham, J. A., & Mesibov, G. Verbal labelling, rehearsal, and short-term memory. *Cognitive Psychology*, 1970, **1**, 47–58.

Hall, J. F. *Verbal learning and retention*. Philadelphia: Lippincott, 1971.

Hintzman, D. L. Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1974.

Hull, C. L. *Principles of behavior: An introduction to behavior theory*. New York: Appleton-Century-Crofts, 1943.

Hull, C. L. The place of innate individual and species differences in a natural science theory of behavior. *Psychological Review*, 1945, **52**, 55–60.

Ingalls, R. P., & Dickerson, D. J. Development of hypothesis behavior in human concept learning. *Developmental Psychology*, 1969, **1**, 707–716.

Jablonski, E. M. Free recall in children, *Psychological Bulletin*, 1974, **81**, 522–539.

Jenkins, J. J. Can we have a theory of meaningful memory? In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1973.

Keeney, T. J., Cannizzo, S. R., & Flavell, J. H. Spontaneous and induced verbal rehearsal in a recall task. *Child Development*, 1967, **38**, 953–966.

Kendler, H. H., & Kendler, T. S. Vertical and horizontal processes in problem solving. *Psychological Review*, 1962, **69**, 1–16.

Keppel, G. Verbal learning in children. *Psychological Bulletin*, 1964, **61**, 63–80.

Kingsley, P. R., & Hagen, J. W. Induced vs. spontaneous rehearsal in short-term memory in nursery school children. *Developmental Psychology*, 1969, **1**, 40–46.

Kuenne, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *Journal of Experimental Psychology*, 1946, **36**, 471–490.

Levine, M. Hypothesis behavior by humans during discrimination learning. *Journal of Experimental Psychology*, 1966, **71**, 331–338.

Lewin, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.

Mackintosh, N. J. Overtraining, extinction, and reversal in rats and chicks. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 31–36.

Mackintosh, N. J. Further analysis of the overtraining reversal effect. *Journal of Comparative and Physiological Psychology, Monograph Supplement*, 1969, **67**, 1–18.

Mackintosh, N. J. Attention and probability learning. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts, 1970.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation*. Vol. 1. New York: Academic Press, 1966.

Mandler, G., & Stephens, D. The development of free and constrained conceptualization and subsequent verbal memory. *Journal of Experimental Child Psychology*, 1967, **5**, 86–93.

Medin, D. L. Subproblem analysis of discrimination shift learning. *Behavior Research Methods and Instrumentation*, 1973, **5**, 332–336.

Peterson, L. R., Wampler, R., Kirkpatrick, M. G., & Saltzman, D. Effect of spacing presentations on retention of a paired-associate over short intervals. *Journal of Experimental Psychology*, 1963, **66**, 206–209.

Restle, F. Discussion: Section I. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1974.

Richard, J. F., Cauzinille, E., & Mathieu, J. Logical and memory processes in a unidimensional concept identification task by children and adults. *Acta Psychologica*, 1973, **37**, 315–331.

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.

Scribner, S. Developmental aspects of categorized recall in a West African society. *Cognitive Psychology*, 1974, **6**, 475–494.

Sharp, D. W., & Cole, M. The influences of educational experience on the development of cognitive abilities manifested in formal tests. *Final report to the National Institute of Education*, 1974.

Skinner, B. F. A case history in scientific method. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959.

Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, 1937, **44**, 430–444.

Sperber, R. D. Developmental changes in effects of spacing of trials in retardate discrimination learning and memory. *Journal of Experimental Psychology*, 1974, **103**, 204–210.

Sperber, R. D., Greenfield, D. B., & House, B. J. A nonmonotonic effect of distribution of trials in retardate learning and memory. *Journal of Experimental Psychology*, 1973, **99**, 186–198.

Stevenson, H. W. *Children's learning*. New York: Appleton-Century-Crofts, 1972.

Sutherland, N. S., & Mackintosh, N. J. *Mechanisms of animal discrimination learning*. New York: Academic Press, 1971.

Tighe, T. J. Subproblem analysis of discrimination learning. In G. H. Bower (Ed.), *Psychology of learning and motivation*. Vol. 7. New York: Academic Press, 1973.

Tighe, T. J., Glick, J., & Cole, M. Subproblem analysis of discrimination shift learning. *Psychonomic Science*, 1971, **24**, 159–160.

Trabasso, T., & Bower, G. H. *Attention in learning*. New York: Wiley, 1968.

Tulving, E. Subjective organization in free recall of "unrelated words." *Psychological Review*, 1962, **69**, 344–354.

Tulving, E., & Donaldson, W. *Organization of memory*. New York: Academic Press, 1972.

Tulving, E., & Osler, S. Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, 1968, **77**, 593–601.

Tulving, E., & Pearlstone, Z. Availability vs. accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 381–391.

Underwood, B. J. Ten years of massed practice on distributed practice. *Psychological Review*, 1961, **68**, 229–247.

Wagner, D. A. The development of short-term and incidental memory: A cross-cultural study. *Child Development*, 1974, **45**, 389–396.

Werner, H. *Comparative psychology of mental development*. New York: International Universities Press, 1948.

Werner, H. The concept of development from a comparative and organismic point of view. In D. B. Harris (Ed.), *The concept of development*. Minneapolis: University of Minnesota Press, 1957.

Zeaman, D., & House, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963.

Zeaman, D., & House, B. J. Interpretations of developmental trends in discriminative transfer effects. In A. D. Pick (Ed.), *Minnesota symposia on child psychology*. Vol. 8. Minneapolis: University of Minnesota Press, 1974.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

5

From Classical Conditioning to Discrimination Learning

N. J. Mackintosh

University of Sussex

I. INTRODUCTION

It is traditional to introduce a discussion of discrimination learning with the assertion that discrimination is of central importance in psychology. It concerns, one is told, such fundamental topics as choice and preference; without discrimination we could not learn to respond appropriately to different situations; without discrimination, we could not tell a hawk from a handsaw, or claret from burgundy.

Discrimination learning may well be important, but more because it pervades most situations in which learning is studied than because of the exciting sound of terms like choice or preference. Neither of these processes appears to be involved in many of the situations in which discrimination learning is studied. Subjects make few if any choices in differential classical conditioning, and may have no preference between the schedules of reinforcement correlated with different stimuli in a free-operant discrimination, although showing, by differences in their patterns of responding, that they have discriminated between them. Indeed, discrimination learning has been studied in such a variety of ways that it is something of a problem to find a single definition adequate to encompass all possible paradigms. What is there in common between differential classical conditioning, a free-operant multiple schedule, a simultaneous discrimination in a T maze, and delayed matching to sample? Little more than this—in each, the subject is exposed to two or more stimulus situations; these situations are correlated with different schedules of reinforcement; and the experimenter infers

that a discrimination has been learned from the emergence of differences in the subject's behavior in the two situations. A definition as wide as this, however, is relevant to virtually all situations in which learning is studied. In a simple experiment on classical conditioning, a reinforcer occurs in the presence of one set of stimuli [shortly after the onset of the conditioned stimulus (CS)], which alternates with a different set of stimuli (the experimental situation in the absence of the CS), which is correlated with no reinforcement. To measure conditioning, the experimenter records differences in the subject's behavior in the presence and absence of the CS. In most free-operant experiments, the subject spends an entire session in the experimental situation, but the probability of a response being reinforced is rarely constant throughout the session. Even if the experimenter does not arrange any explicit discriminative stimuli, the availability of reinforcement is often signaled by the passage of time, or the occurrence of a previous reinforcer, and the subject's behavior usually reveals that this discrimination is learned. Discrimination learning therefore is of basic importance because the study of classical or instrumental conditioning requires the arrangement of correlations between events in the subject's environment, and learning these correlations requires the learning of certain discriminations.

Although it is not unusual to find a statement of its fundamental importance at the outset of a chapter on discrimination learning, the reader usually finds himself immediately thereafter subjected to a detailed description of typical procedures for studying discrimination learning, and immersed in theoretical and experimental problems apparently unique to those procedures. Although it is obvious that no account of discrimination learning can ignore all procedural details, I hope it will be possible in what follows to show the relationship between some of the more important problems of discrimination learning, and those of supposedly simpler situations.

II. CLASSIFICATION OF PROCEDURES

Although no simple classificatory scheme will be entirely satisfactory, one can distinguish between two main procedures used in experiments on discrimination learning. In one, the discriminative stimuli are presented successively, and the experimenter typically measures the strength of a single response (whether or not the subject salivates, blinks, or pecks a response key); in the other, the discriminative stimuli are presented simultaneously, and the experimenter records the subject's choices between them. In a typical successive discrimination, there are two discriminative stimuli, one correlated with reinforcement, the other not, presented in a quasi-random

alternating sequence. Examples include differential classical conditioning, in which CS+, signaling the unconditioned stimulus (US), alternates with CS— signaling its absence; free-operant multiple schedules, in which a pigeon's key pecks are reinforced according to a particular schedule in the presence of one stimulus (S+), and are not reinforced in the presence of another (S—); and discrete-trial go–no go discriminations in which a rat is placed on alternate trials in either a black or a white runway, with food available in the goal box of the black, but not of the white, runway.

In a typical simultaneous discrimination, the subject is exposed to two alternatives on each trial, with choice of one being reinforced, and the other not. The simplest example is a simultaneous black–white discrimination in a T maze, discrimination box, or jumping stand, where the black and white alleys, goal boxes, or doors appear side by side on each trial, with black sometimes on the left and sometimes on the right, and choice of black is reinforced and choice of white is not. Simultaneous-choice procedures have also been used, however, in free-operant situations (concurrent schedules), and to study more complex discriminations, for example, problems often described as conditional discriminations. A pigeon given a choice between response keys illuminated with red or green light on each trial might be required to learn that which alternative was correct depended on the color of a third, sample key on that trial.

Neither the value nor the accuracy of these definitions should be accepted on trust. One version of what would here be called a simultaneous discrimination might on other criteria be called a successive discrimination: in a so-called successive-conditional problem, a subject might be exposed to two black alternatives or two white alternatives on each trial. When the stimuli are black, choice of the right-hand alternative is reinforced; when they are white, a left choice is reinforced. This can be described either as a simultaneous discrimination between black on the left and black on the right on some trials, and white on the left and white on the right on others, or as a successive discrimination in which black and white are the successively presented discriminative stimuli, with one signaling that a response to the left is reinforced, and the other that a response to the right is reinforced.

The important point is that, here and elsewhere, these definitions are not entirely arbitrary: they have empirical content, and experimental evidence can be brought to bear on the question of which definition or description is correct. If the first problem for discrimination learning theory, therefore, is to provide an adequate description of the various situations used, this amount to discovering what are the effective stimuli that control the subject's behavior in any situation.

III. CONDITIONING-EXTINCTION THEORY

A. Analysis of Successive Discriminations

There seems to be less problem, and there has certainly been less dispute, about the proper definition of the effective stimuli in successive discriminations than in simultaneous discriminations. A reasonable strategy therefore will be to consider first a particular theoretical account of successive discriminations. We can then see how this account has been extended to cover the case of simultaneous discriminations, and whether this extension is successful. The account originated with Pavlov (1927), but was more systematically developed by Spence (1936, 1937) and Hull (1952). It is usually known as conditioning-extinction theory.

In simple cases of differential conditioning or go–no go discriminations, the experimenter presents two stimuli in random sequence (a metronome and a buzzer, or two tones of different frequency, or a red and a green key light). The presentation of one stimulus is correlated with the availability of reinforcement, whereas the other is not, and successful discrimination learning is inferred from the observation that the reinforced stimulus comes to elicit a conditioned response (CR) or control the occurrence of an appropriate instrumental response, whereas the nonreinforced stimulus controls a significantly lower level of responding. It seems simplest to assume that the effective stimuli are those defined by the experimenter—tones of particular amplitudes and frequencies, lights of particular hue and brightness—one of which is associated with reinforcement, the other with nonreinforcement.

With the aid of some data, we can be more specific. Figure 1 shows the course of conditioning of the nictitating membrane response in three groups of rabbits (Moore, 1972). Group T+ received only reinforced trials to a tone (actually, for purposes of control, to two tones; the data shown are for performance to one of the tones); Group T+N− received reinforced trials to a tone alternated with nonreinforced trials to a noise; and Group T+L− received reinforced trials to the tone alternated with nonreinforced trials to a light. With continued training, there is in all three groups a steady increase in the probability of a response to CS+. We may assume that reinforcement of CS+ results in the development of excitatory conditioning, it being understood that the term excitation signifies no more than the ability of a stimulus to elicit a CR or control instrumental responding as a consequence of its correlation with reinforcement. By the principle of generalization, if excitation is conditioned to CS+, other stimuli will acquire excitation in proportion to their similarity to CS+. One might expect to see, therefore, generalized responding appearing to CS−, with the probability of such responding depending on the similarity of CS− and

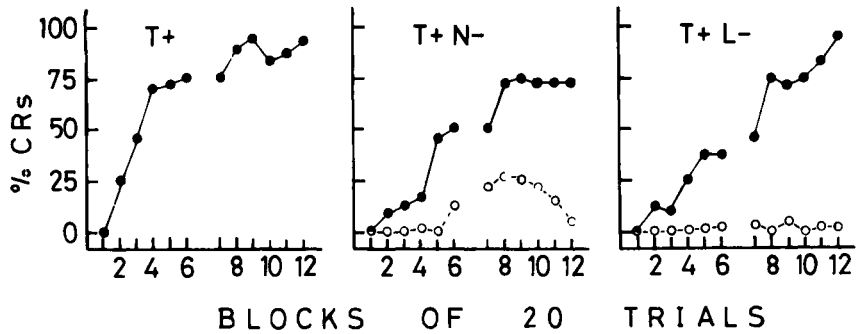

FIGURE 1 Differential conditioning in rabbits: Open circles, S—; filled circles, S+. For Group T+, all trials were reinforced (in fact reinforced trials with a different tone were interspersed with those shown here.) For Groups T+N— and T+L—, reinforced trials to the tone were interspersed with nonreinforced trials to a noise and to a light respectively. (After Moore, 1972.)

CS+. Since CS— is never reinforced, however, it will acquire no excitation of its own, and if it is sufficiently distinct from CS+, there will be no evidence of responding to CS—. The initial phase of discrimination training shown in Figure 1 illustrates all of these points: Group T+N— shows a significant increase in CRs on CS— trials, whereas in Group T+L— for whom CS+ and CS— are from different modalities, there is essentially no evidence of generalized responding to CS—.

If responding is elicited by CS— through generalized excitation, it should extinguish with continued nonreinforcement, just as responding to a former CS+ may be extinguished by nonreinforcement. Group T+N— shows an appropriate decline in responding to CS—. Without entering into a prolonged discussion on the nature of excitation and inhibition, we may refer to the effects of nonreinforcement as resulting in inhibitory conditioning, and assume that inhibition conditioned to a particular stimulus will generalize to other stimuli in proportion to their similarity, in much the same way as excitatory conditioning generalizes. In the case in which CS+ and CS— are relatively similar therefore, one might expect to find that the level of responding maintained by CS+ is less than that observed either when there is no CS— or when CS+ and CS— are highly dissimilar. There is, indeed, some suggestion that Group T+N— respond less to CS+ than do either of the other groups.

B. Analysis of Simultaneous Discriminations

An account relying on the simple notions of excitatory and inhibitory conditioning and generalization therefore seems adequate to handle these admittedly quite straightforward data. More impressive was Spence's extension of the theory to the case of simple simultaneous discriminations.

Spence (1936, 1937) argued that in the standard simultaneous black–white discrimination, the situation must be conceptualized as consisting of four component stimuli: black, white, left, and right. Discrimination training would increase the excitatory or inhibitory properties of each of these component stimuli in accordance with the schedule of reinforcement or nonreinforcement associated with each. Generalization between black and white, and between left and right, would also have to be taken into consideration in calculating the final associative strength of each stimulus. Finally, Spence proposed the following simple rule for predicting the subject's choice on each trial. On trials in which black is on the left and white on the right, the associative strengths of black and left are combined, and compared with the combined strengths of white and right; if they are greater, the subject chooses black on the left, if smaller, white on the right. On trials in which black is on the right and white on the left, then the associative strengths of black and right are compared with those of white and left, and choice is determined accordingly. By making certain quantitative assumptions about the rates of change of excitation and inhibition, and the initial strengths of the component stimuli, Spence was able to predict with considerable success a number of details of the rat's performance on such discriminations.

C. Reactions to Conditioning-Extinction Theory

Much of the history of research and theorizing on discrimination learning may be understood as a reaction to this conditioning-extinction theory. From a variety of viewpoints, the argument has been advanced that the processes involved in discrimination learning are somehow *sui generis,* emerging from the complex demands placed on the subject by discriminative experiments, and not to be reduced to the trivial, associative processes of simple conditioning experiments. This is an important argument, and although I shall suggest that it has often been advanced in an inappropriate form, there can be little question but that properly understood it should help to provide a satisfactory theory of discrimination learning. In the meantime, however, we may first consider whether the account thus far considered provides an adequate definition of the effective stimuli in successive and simultaneous discriminations.

1. Situational Stimuli in Successive Discriminations

Psychological experiments are not conducted in a vacuum. Conditional or discriminative stimuli are always presented in a particular experimental situation or context.[1] The situation or context comprises a particular set

[1] As is noted below (Section IV.B.2), there is often nothing in the schedule of reinforcement to distinguish between what the experimenter chooses to call discriminative and contextual stimuli.

of stimuli, and it would be surprising if such stimuli in no way entered into association with reinforcement during the course of conditioning or discrimination training. An adequate account of simple discrimination learning, therefore, will almost certainly have to make reference to contextual or situational stimuli.

For purposes of initial analysis, we may represent the population of contextual stimuli by the term X. An experiment on differential conditioning between two auditory stimuli, A_1+ and A_2-, therefore, should properly be regarded as a sequence of three types of trial. A_1X, A_2X, and X, of which only the first is correlated with reinforcement. Similarly, an experiment on simple conditioning with a single auditory CS should be regarded as a sequence of two types of trials, $AX+$ alternating with $X-$. Several implications of this manner of looking at simple learning situations are considered below (Sections IV.B and C). For the moment we may concentrate on one point, related to our earlier discussion of the effect of the similarity of S+ and S− on the rate of discrimination learning. If the occurrence of a reinforced trial in a simple conditioning experiment is represented by the addition of A to the contextual stimuli X, the discriminability of trials from the intertrial interval will presumably be a function of the intensity of A. With an intense CS, therefore, there will be little generalization of excitation from reinforced AX trials to the nonreinforced intertrial interval X, and similarly little generalization of inhibition from X to AX. With a less intense CS, however, and greater generalization between AX and X, conditioning will progress more slowly, both because of greater generalization of inhibition from X to AX, and because of a greater generalization of excitation from AX to X. That the rate of conditioning is a function of the intensity of the CS formed part of the Pavlovian "law of strength," and was described by Hull (1949) as the phenomenon of "stimulus intensity dynamism." Logan (1954) and Perkins (1953) independently argued that the effects of stimulus intensity on conditioning did not require any new construct, as Hull supposed, but were a straightforward consequence of generalization between the CS and the background in the manner outlined. Although other factors may indeed contribute to the overall effects of stimulus intensity (e.g., Grice, 1968), there can be little doubt that this analysis points to an important contributing factor (Gray, 1965; Kamin, 1965).

2. *Component, Compound, and Configuration in Simultaneous Discriminations*

Taking contextual stimuli into consideration increases the power of the account provided by conditioning-extinction theory of successive discriminations and simple conditioning. Objections to the theory, however, have

usually concentrated on its extension to simultaneous discriminations, and have been directed, therefore, largely against Spence's analysis of such discriminations. It has frequently been argued that the stimulus situation of a simultaneous discrimination is more complex than is allowed by Spence. Certainly it cannot be disputed that there are several plausible descriptions of the effective stimuli of a simultaneous black–white discrimination other than that proposed by Spence. A brief enumeration would include:

a. Component theory (Spence). Four stimuli, all of which are present on each trial: black, white, left, and right.

b. Compound theory (e.g., Medin, in press; Spiker, 1970). Four compound stimuli, two of which are present on each trial: black on left and white on right on some trials; white on left and black on right on others.

c. Configuration theory (e.g., Gulliksen & Wolfle, 1938). Two configurations, one of which is present on each trial: black to left of white on some trials; white to left of black on others.

d. Relational theory (e.g., Köhler, 1918). Two stimuli, both of which are present on each trial: one is defined as darker than the other; the second is defined as brighter.

Elaborate as it may seem, it is doubtful whether even this list is exhaustive. All of these analyses, for example, and in particular the third and fourth, assume that the subject inspects both alternatives on every trial. There is, in fact, some reason to question this assumption. There is evidence that subjects who develop a consistent position habit during the course of training on a simultaneous discrimination (that is, respond consistently to one side), are actually treating the problem as though it were a successive go–no go discrimination between the black and white stimuli appearing in that position (Hall, 1974a; Mahut, 1954).

Considerable research effort has been expended in attempts to decide between these different descriptions, and this is discussed in some detail by Medin and Cole (Chapter 4 of this volume). For the present we may briefly touch on two points. Spence himself recognized that there were problems with a component analysis of all simultaneous discriminations (Spence, 1952). It predicts, for example, that a successive conditional discrimination (with two black stimuli, left is correct; with two white stimuli, right is correct) will be insoluble, for each of the four components is equally often reinforced and not reinforced. Spence allowed that under such circumstances subjects could learn to respond to compounds, thus learning that black on left and white on right were reinforced, and the other two compounds were not. An obvious problem with this analysis, however, is that it implies some rather complex and ill-defined processes whereby subjects, recognizing that one mode of analysis does not work, learn to shift

to another. Whether this solution must be accepted is not clear: it may be preferable to assume that all simultaneous discriminations are always solved as compounds.

3. *Transposition*

A question that has been more persistently raised against Spence's analysis of simultaneous discriminations, is whether the effective stimuli should not be described in relational terms. According to a long line of theorists, a subject solving a simultaneous black–white discrimination, does not respond to the absolute properties of the black and white stimuli (for example, their reflectances), but to the relationship between them. Köhler's (1918) demonstrations of transposition appeared to provide strong support for this account: a subject trained to select the darker of two stimuli in a simultaneous brightness discrimination would select one even darker in preference to the original S+ on a subsequent test trial.

In one of the more elegant early demonstrations of the virtues of quantitative theorising, Spence (1937) showed that appropriate assumptions about the shapes of excitatory and inhibitory generalization gradients permitted the prediction of transposition from simple conditioning-extinction theory. The derivation is illustrated in Figure 2. Reinforcement of a dark

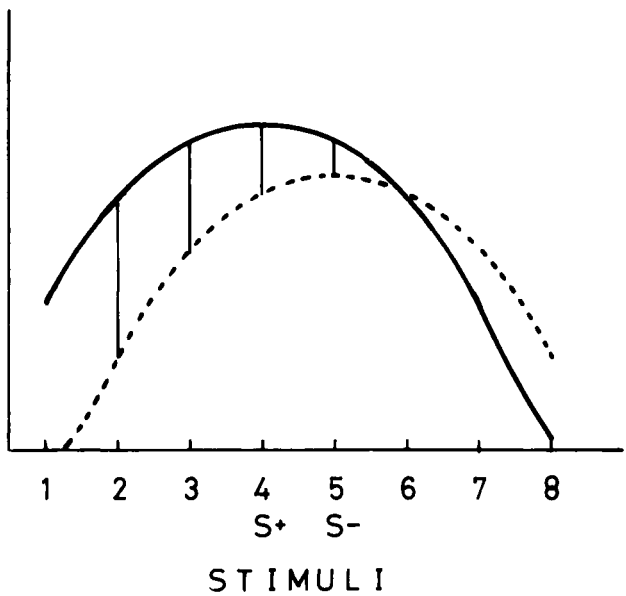

FIGURE 2 Hypothetical gradients of excitation (solid line) and inhibition (dashed line) formed after reinforcement at Stimulus 4 and nonreinforcement at Stimulus 5. (After Spence, 1937.)

gray stimulus (Number 4 in the example shown) produces excitatory conditioning to that stimulus, which generalizes to other stimuli along the brightness dimension in the manner shown. Nonreinforcement of a lighter gray (Number 5) produces inhibitory conditioning which also generalizes to other stimuli. The difference between these two gradients at each point on the dimension defines the net associative strength of each of the stimuli, 1 to 8. Choice of one stimulus over another is determined, other things being equal, by the relative values of their associative strengths, and, as can be seen, in a test between 2 and 4 or 3 and 4, the subject will choose the darker stimulus in preference to S+.

Spence's account of transposition received progressive elaboration in the 25 years following its publication, culminating in the work of Hanson (1959) and Honig (1962), who used free-operant procedures to study gradients of wavelength generalization in pigeons following discrimination training between various wavelengths. Hanson reinforced pigeons for pecking a key illuminated with light of 550 nm. For a control group there was no S—, but for different discrimination groups, S— varied from 555 to 590 nm. The generalization test consisted of the presentation of stimuli ranging from 480 to 620 nm, and the gradients produced by several of Hanson's groups are shown in Figure 3. The control group responded more frequently to S+ (550 nm) than to any other stimulus, but the discrimination groups all responded more rapidly to wavelengths shorter than S+ than to S+ itself. Honig's data appeared to confirm that this shift in the peak of a postdiscrimination gradient exactly paralleled the transposition shown in choice tests, and thus confirmed the central assumption of Spence's analysis.

A relational view of transposition suggests that subjects given the opportunity to do so should compare discriminative stimuli with one another, and respond to the relationship between them. Transposition would therefore depend on the simultaneous presentation of the discriminative stimuli. Hanson's results showed that the analogue of transposition occurred in successive discriminations. It seemed, therefore, that there was no need to distinguish between the description of the stimuli in simultaneous and successive discriminations; conditioning-extinction theory explained the results of both procedures without recourse to a relational interpretation.

Although at first sight impressive, there can be little doubt that this account is inadequate. Hanson's peak shift may indeed be regarded as confirmation that excitatory and inhibitory gradients interact in very much the way that Spence suggested. It is not, however, the perfect analogue of transposition. Transposition does, in fact, occur more readily in simultaneous than in successive discriminations, just as relational theory predicts (Riley, Ring, & Thomas, 1960; Sachs, 1969), and it does depend on the opportunity to compare the discriminative stimuli (Riley, Goggin, &

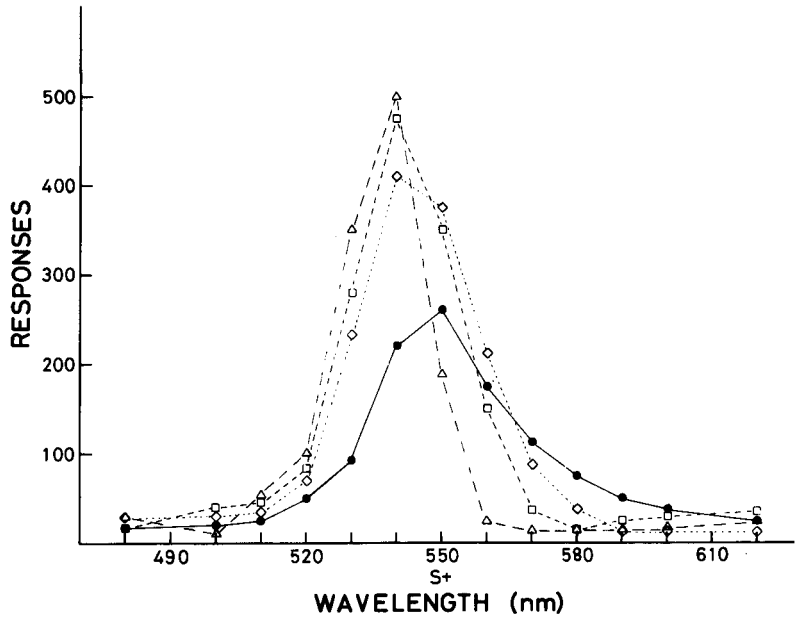

FIGURE 3 Generalization gradients to other wavelengths in pigeons, after reinforcement for responding to 550 nm. The control group (filled circles) received only reinforced trials; for the remaining groups, reinforced trials to 550 nm were interspersed with nonreinforced trials to longer wavelengths: (◇)S–, 590 nm; (□)S–, 570 nm; (△)S–, 560 nm. (After Hanson, 1959.)

Wright, 1963); moreover, subjects given simultaneous discrimination training will show transposition to a new pair of stimuli on a choice test, when a successively presented generalization test reveals no evidence of a peak shift to those stimuli (Rudolph, 1967, cited by Riley, 1968; Sachs, 1969).

D. Conclusion

An adequate account of discrimination learning will require a more elaborate, and possibly a more flexible, description of the effective stimuli than that sufficient for most instances of simple conditioning. In view of the variety and complexity of the stimuli used in discriminative experiments, this conclusion is not entirely surprising, nor does it really strike at the heart of conditioning-extinction theory. We can surely allow that the relationship between two stimuli can itself act as a stimulus, which may then acquire excitatory or inhibitory properties as a consequence of its correlation with reinforcement, without thereby rejecting the central assumptions of the theory. We are required only to extend the definition of the stimuli which may affect our subjects.

It is time, however, to turn to arguments that have been thought to point to differences in the basic associative processes involved in complex discrimination learning and in simple conditioning. One unfortunate consequence of the belief that discrimination learning cannot be reduced to simpler processes of conditioning has been the tendency, remarked earlier, to treat the problems and theories of discrimination in isolation. But if the importance of discrimination learning lies in its pervasiveness, and if discriminative processes must be assumed to operate in even the simplest cases of associative conditioning, then one might do well to turn the argument on its head: an adequate account of conditioning, one might suppose, must itself be complex enought to incorporate these complex processes. Although the phenomena of discrimination learning may require a more elaborate analysis than that envisaged by conditioning-extinction theory, they are in this respect no different from those of conditioning itself. It is our theories of conditioning that need elaboration, and once we have provided this, we shall find it easier to see the continuity between conditioning and discrimination learning.

The argument can best be illustrated by analysis of a particular case history. A good example is the occurrence of contrast effects in discrimination learning. The data shown in Figure 1 provided good evidence of generalization between S+ and S— during the course of discrimination learning. As conditioning-extinction theory predicts, generalization of inhibition from S— to S+ depressed performance to S+, whereas generalization of excitation from S+ to S— enhanced performance to S—. Unfortunately, these data are very far from typical. If an experimenter programs different schedules of reinforcement in the presence of different discriminative stimuli, he will frequently observe an exaggeration in the pattern of behavior that would have been maintained by each of these schedules in isolation. Rate of responding to a stimulus signaling a moderate probability of reinforcement may be increased by introducing trials with a second stimulus signaling nonreinforcement: this is a defining example of positive behavioral contrast (Reynolds, 1961). Conversely, responding to a stimulus signalling small or infrequent reinforcement may be depressed by concurrent exposure to a stimulus signaling larger or more frequent reinforcement: this is a defining example of simultaneous negative contrast (Black, 1968). Interactions such as these are, on the face of it, the exact opposite of what one would be led to expect from a theory in which the generalization of excitatory and inhibitory tendencies plays a central role.

The problem posed by the occurrence of contrast effects, however, is not that such effects are unique to studies of instrumental discrimination learning, for Pavlov observed what appear to be analogous effects in experiments on differential conditioning, and labeled them positive and negative

induction. The problem, it should be clear, arises only for a particular *theory* of discrimination learning. If contrast effects force some rethinking of simple theories of discrimination learning, this is no more than the rethinking necessitated for theories of conditioning by the phenomena of induction. Indeed, a persuasive account of the positive behavioral contrast observed in free-operant multiple schedules with pigeons is that it is a direct consequence of a process of classical conditioning.

In a typical experiment on behavioral contrast, a pigeon's response key is alternately illuminated with red and green light (S_1 and S_2), and the pigeon's key pecks are initially reinforced on identical variable-interval schedules in the presence of both stimuli. After a number of sessions, reinforcement is omitted in the presence of S_2. Appropriately enough, responding to S_2 extinguishes, but although the schedule associated with S_1 is unchanged, the pigeon responds much more rapidly to S_1 than before. Now one consequence of the decline in the probability of reinforcement in S_2 is an increase in what may be termed the informativeness of S_1, which is now the sole predictor of reinforcement in the experimental situation. There is evidence, discussed by Rudy and Wagner (Chapter 7, Volume 2 of this *Handbook*) and in further detail below (Section IV.C.1) that the effectiveness of a CS may depend not only on the probability of reinforcement in its presence, but also on the extent to which it is the best available predictor of reinforcement. The associative strength of a stimulus, and hence its capacity to elicit CRs, may depend more on its relative validity compared to that of other stimuli in the situation (Wagner, 1969a), than on its absolute validity. The omission of reinforcement in S_2 increases the relative validity of S_1, and the rate of pecking in S_1 may increase because pecking, as the data on autoshaping suggest (Brown & Jenkins, 1968; see Hearst, Chapter 5 in Volume 2) is a response classically elicited in pigeons by stimuli signaling the availability of food (Boakes, Halliday & Poli, 1975; Hearst & Jenkins, 1974). So far from being inconsistent with a conditioning model of discrimination learning, therefore, behavioral contrast may well depend on the use of a supposedly instrumental response, key pecking in pigeons, that is unusually susceptible to classical contingencies. Certainly, such contrast effects do not occur when pigeons are trained to press a treadle rather than peck a key (Hemmes, 1973; Westbrook, 1973), and are at best unreliable when rats are trained to press a lever (Pear & Wilkie, 1971).

The problem therefore is to explain why the classical pairing of a localized visual stimulus with an appetitive reinforcer should result in the direction of pecking responses to the visual stimulus, and why changes in the relative validity of a stimulus, without any change in the absolute schedule of reinforcement associated with it, should increase its ability to control classically conditioned responses. This is not the place to consider the first

of these problems; it is to be hoped that the second will seem less problematical as we proceed.

The criticisms of conditioning-extinction theory of discrimination learning advanced by gestalt or cognitive theorists have typically been based on the assertion that it is impossible to reduce the richness and complexity of discrimination learning to the conditioning and generalization of excitatory and inhibitory processes, as these are revealed in studies of simple conditioning. My argument is that it may be possible to bridge the gap between opposing theorists by showing that the processes of excitatory and inhibitory conditioning are themselves notably more complex than has sometimes been supposed. An adequate theory of simple conditioning will itself be sufficiently subtle and powerful to account for the major phenomena of discrimination learning. The fundamental associative processes of discrimination learning are the same as those engaged in experiments on classical conditioning, and experiments on discrimination learning serve only to throw into sharper relief some of the central issues in the analysis of classical conditioning. In the remainder of this chapter, I shall consider two main topics, those of selective association and transfer of training, in an attempt to see how far this synthesis may be carried.

IV. SELECTIVE ASSOCIATION

If animals learn to approach and avoid particular discriminative stimuli because they learn that such stimuli signal reinforcement and nonreinforcement, then perhaps the central question for theories of discrimination learning, as for theories of conditioning, is how stimuli become established as signals for reinforcement. Traditional associationist answers stress the importance of close temporal contiguity with reinforcement. In stimulus–response theory, therefore, if a response repeatedly occurs in the presence of a stimulus and is consistently followed by a reinforcing event, this should be sufficient to establish a strong tendency for that response to occur again whenever that stimulus is presented.

A. Competition between Stimuli: Overshadowing

Experiments on discrimination learning and generalization have long provided numerous apparent exceptions to this general proposition. A traditional solution has been for conditioning-extinction theorists to borrow a leaf from their opponents' notebook, and to argue that the nature of the stimuli in many discriminative experiments is such that important new principles must be invoked. Thus, Spence (1936, 1937) qualified his account of discrimination learning by allowing that the adequate perception of localized discriminative stimuli might require the establishment of appro-

priate receptor-orienting responses. The reason, therefore, why a particular stimulus does not acquire significant control over responding may be because the subject has not yet learnt to look at it. The simpler, diffuse stimuli, such as lights and tones, typically used in experiments on classical conditioning do not require such elaborate orientation for their adequate perception.

Although the role of observing responses in instances of selective association has been more frequently assumed than experimentally demonstrated, there is little question but that monkeys, for example, tend to look preferentially at the visual stimuli which they touch with their fingers (Stollnitz, 1965), whereas pigeons learn more about visual stimuli they actually peck at than those located some distance from the point of pecking (Jenkins, 1973). Nevertheless, it is by now abundantly clear that selective association occurs with diffuse stimuli in experiments on classical conditioning just as readily as with discrete stimuli in experiments on discrimination learning. As Rudy and Wagner (Chapter 7, Volume 2) discuss, theories of conditioning must confront this ubiquitous phenomenon: the associative strength of a particular stimulus does not depend solely on its own salience and history of reinforcement, but also on the salience and reinforcement histories of all other stimuli (including situational stimuli) with which it is presented. As a first approximation, one can assert that there appears to be some competition between stimuli for the acquisition of associative strength or signaling properties.

Although such interactions between stimuli were first noted by Pavlov (1927) in his observation of the phenomenon of overshadowing, the subsequent investigation of selective association was for a long time confined to experiments on instrumental learning (Lashley, 1942; Mackintosh, 1965a). It is only recently that the work of Kamin (1968, 1969) on blocking and overshadowing in conditioned suppression, and of Wagner on relative validity effects in conditioned suppression and eyelid conditioning (Wagner, 1969a, b; Wagner, Logan, Haberlandt, & Price, 1968) has reasserted the central importance of such selectivity in classical conditioning. One good reason for this concentration on conditioning, indeed, has been that the use of diffuse stimuli, whose exposure is directly controlled by the experimenter rather than by the subject, may not only serve to eliminate the type of explanation favored by Spence, but also permit better control of certain factors typically left uncontrolled in instrumental experiments. As Wagner (1969b) has pointed out, it may be difficult to equate the actual schedule of reinforcement received by experimental and control subjects in an instrumental experiment on blocking or overshadowing. In an instrumental study of blocking, for example, pigeons might be trained on a compound color and line-tilt discrimination (with S+ a vertical line on red, and S− a horizontal line on green); the control group receives no prior training, whereas the experimental group is pretrained on the red–green

color discrimination. When tested for control by the line, the experimental subjects show that they have learned less than the controls: pretraining on color has apparently blocked learning about the lines. But it is obvious that the two groups could not have had the same experience with the lines; although the control subjects may have made a large number of errors during compound training, the experimental group, if their pretraining was effective, would have continued to respond only to the reinforced color, and would thus make no errors. Might not any difference in control by the lines be a consequence of this difference in errors? Although it is possible to design instrumental experiements that are not obviously open to this objection (e.g., Mackintosh, 1965b; Mackintosh & Honig, 1970), the fact remains that many have been.

It may be no accident, therefore, that most recent research on selective association has employed classical conditioning paradigms, and that, whereas a decade ago, theoretical analysis of selective association was provided largely by models of instrumental discrimination learning (Lovejoy, 1968; Sutherland, 1964; Zeaman & House, 1963), several of which could not readily be applied to the case of classical conditioning at all, what is now probably the most widely accepted account of such effects is a theory proposed by Rescorla and Wagner (1972) expressly designated as a theory of classical conditioning.

B. Irrelevant Stimuli in Generalization and Discrimination

1. Ubiquity of Irrelevant Stimuli

Some of the reasons for the earlier concentration on discrimination learning as the paradigm for studying selective effects are doubtless those of historical accident (for example, Lashley's own interest in discrimination learning, and his mistrust of conditioning procedures). But there are also features inherent in the design of experiments on discrimination learning and generalization that were always likely to have brought this issue to the fore. A simultaneous visual discrimination necessarily contains both relevant and irrelevant stimuli: on any one trial, the alternatives between which the subject is required to choose differ not only in their visual characteristics, but also in their spatial location. It is, moreover, obvious to even the most casual investigation that rats trained on such a visual discrimination are not indifferent to this irrelevant aspect of the problem: before solving the problem, they will typically respond systematically to one or other position. An adequate theory of simultaneous discrimination learning therefore must come to grips with the problems posed by the development of these position habits during the course of training, despite the equal schedules of reinforcement associated with the two positions, and the eventual superseding of these position habits by the correct solution, despite

the fact that the schedules of reinforcement associated with relevant and irrelevant stimuli do not change during the course of the experiment. It is surely no accident that noncontinuity theory of discrimination learning, stressing the possibility of selective learning and control, was born of Lashley's and Krechevsky's observations on the development of position habits in rats trained on simultaneous visual discriminations.

Although the occurrence of systematic position habits in simultaneous visual discriminations may draw attention to the importance of irrelevant stimuli in any theoretical analysis of such learning, we have already remarked that any experimental situation, including the simplest nondiscriminative study of classical conditioning, includes numerous stimuli other than those in which the experimenter is specifically interested (Section III.C.1). Both in simple and in differential conditioning (or successive discrimination learning), CSs are presented against a background of situational stimuli. This was the basis for the prediction of the effects of CS intensity on rate of conditioning in terms of generalization between background and CS; and situational stimuli also play a crucial role in the derivation of a number of predictions in Rescorla and Wagner's (1972) model of Pavlovian conditioning. Once again, therefore although certain issues may be accentuated by typical discriminative experiments, there is little reason to doubt their importance in supposedly simpler situations.

2. *Irrelevant Stimuli Mask the True Slope of Generalization Gradients*

What then are the effects of irrelevant or incidental stimuli in conditioning or discrimination learning? The point is perhaps raised most clearly in experiments on generalization. In a typical operant-generalization experiment, a pigeon is trained to peck a key onto which is projected a vertical white line. Once a steady rate of responding has been established, reinforcement is omitted while the experimenter varies the orientation of the line on the key. Although such a procedure usually results in a reliable decrease in the pigeon's rate of responding as the orientation of the line departs from the vertical (that is, usually produces a reliably sloping gradient of generalization), there are a number of reports of essentially flat gradients, where no evidence of control by the orientation of the line could be detected (e.g., Newman & Baron, 1965). With other stimuli, flat generalization gradients may well be the rule rather than the exception: Jenkins and Harrison (1960), for example, trained pigeons to peck a response key illuminated with white light while a 1000-Hz tone was switched on. Subsequent generalization tests revealed little or no change in rate of responding when the frequency of the tone was varied between 300 and 3500 Hz.

Even if it is well known that the subject is capable of detecting variations along a particular dimension, flat gradients of generalization along that dimension following nondifferential reinforcement should not really be regarded as particularly surprising. The experimenter may describe the experimental situation as one in which a particular stimulus is presented in a given context, and is then varied during a subsequent test period, while the context remains the same. By definition, however, if reinforcement has been scheduled nondifferentially, there is nothing in the experimental situation to single out the experimenter's discriminative stimulus from the context—except the fact that the experimenter plans to vary that stimulus during a subsequent test series. It is hardly surprising that the subject's behavior may remain unaffected when one, possibly trivial, aspect of the total situation is changed, but all other aspects are held constant. If the schedule of reinforcement correlated with the experimenter's discriminative stimulus is no better than that correlated with other contextual stimuli, then both relevant and incidental stimuli may be established as signals for reinforcement. If the latter are more salient, as well as more numerous, then the subject will presumably continue to respond so long as they remain unchanged. Thus, flat gradients of generalization following nondifferential reinforcement may be due to a masking effect. In Jenkins and Harrison's (1960) experiment, for example, it is possible that the pigeons continued to respond as before when the frequency of the tone was changed, not because they were deaf, nor even because they had failed to associate the tone with reinforcement, but because they had also associated other features of the experimental situation, such as the key light, with reinforcement, and these remained unchanged during the generalization test.

This argument was developed by Hull (1952), and there is substantial evidence to support it. Rudolph and Van Houten (in press) repeating Jenkins and Harrison's experiment, trained pigeons in the presence of a 1000-Hz tone to peck a white key in order to obtain food. Like Jenkins and Harrison, they found that subsequent variation in the frequency of the tone had virtually no effect on the pigeons' behavior. But when birds were trained to peck an unilluminated key in a darkened chamber, it was quite possible to obtain reliably sloping gradients when the frequency of the tone was varied. Presumably, therefore, the key light is a more salient stimulus for the pigeons than is the tone: if the light gains control of pecking, its presence on all test trials will serve to maintain a relatively unchanged rate of pecking throughout the test. Only in the absence of this masking stimulus could the control acquired by the tone be demonstrated. There are numerous examples of similar masking effects. One reason why Newman and Baron (1965) observed little control by the orientation of the line projected onto their pigeons' response key, it is now clear, is that the line was displayed on a colored background throughout training and

testing. Explicit comparisons of the control exercised by the orientation of a line shown on a colored or on a black background have shown that the presence of the color may significantly mask the control displayed by the line (Farthing, 1972; Freeman & Thomas, 1967).

Masking effects such as these are surely not very surprising, nor do they pose intractable theoretical problems. The problem is that the unwary experimenter, ignoring the presence of potential masking stimuli in his experimental situation, may attribute a flat gradient to other factors. Heinemann and Rudolph (1963) and Terrace (1966), for example, have argued that the absence of auditory control in Jenkins and Harrison's experiment was due to a lack of implicit differential reinforcement. Their argument, which can be traced back to Lashley and Wade (1946), was that a stimulus will gain control over behavior only if differential reinforcement is explicitly, or implicitly, correlated with variations in that stimulus for the subject. The reason why stimuli projected onto a pigeon's response key rapidly gain control over pecking is that, where reinforcement is contingent on pecking the key, there will be an implicit correlation between the sight of those stimuli and the occurrence of reinforcement. A diffuse stimulus, such as a tone, however, will not vary so much with the occurrence of a pecking response, and will not therefore be differentially correlated with reinforcement.

As we shall see, there is no question but that differential reinforcement, explicitly correlated with the presence and absence of a particular stimulus, may have a marked effect on the control exercised by that stimulus. There is, however, little reason to extrapolate from this observation to the notion that differential reinforcement is a *necessary* condition for the observation of stimulus control. In Rudolph and Van Houten's (in press) experiment, the appearance of reliable auditory control did not depend on altering, implicitly or explicitly, the schedule of reinforcement correlated with the presence or absence of the tone, but on removing a potential masking stimulus. The presence of masking stimuli therefore is a more plausible explanation of the lack of control than is the absence of implicit differential reinforcement.

3. *Irrelevant Stimuli Overshadow the Experimenter's Discriminative Stimuli*

Although Lashley and Wade (1946) were incorrect when they claimed that differential reinforcement was necessary for the establishment of stimulus control, the fact that the claim was unjustified may not invalidate the importance of the theoretical rationale for their argument. In effect, they supposed, a subject might fail to associate one stimulus or aspect of the experimental situation with reinforcement, because he was attending to

some other features, and treated those features as the signal for reinforcement. The experimenter's stimulus might be *overshadowed* by other, presumably more salient, aspects of the experimental situation. It is important to see how this suggestion differs from Hull's. According to Hull (1952), it might be impossible to detect the control acquired by a particular stimulus, because other stimuli also gained control over responding, and masked, during testing, the control acquired by the first. According to Lashley and Wade, a particular stimulus might actually fail to gain control at all, because it was overshadowed, during training, by other stimuli. Masking is a test effect, serving to obscure control that is really there; overshadowing is a training effect, serving to prevent the acquisition of control in the first place.

There is, in fact, little doubt that both processes may operate—and in one and the same experiment. A study by Farthing (1972) illustrates this. Pigeons were trained on a successive discrimination, with reinforced trials signaled by a vertical line on a colored background, and nonreinforced trials signaled by a plain colored key. For one group of birds, the color of the key was the same on positive and negative trials, and the presence or absence of the line therefore served as the only signal for reinforcement. For a second group, however, the color of the key differed on positive and negative trials, and this group therefore had two potential signals for reinforcement. Both groups were given a generalization test to other orientations of the line, shown either on a colored or on a black background. The results of these tests are shown in Figure 4. The presence of the reinforced color during test trials significantly reduced the control displayed by the line in both groups. In addition to this masking effect, however, there was clear evidence that subjects for whom differences in color had been correlated with reinforcement in acquisition showed less control by the line than did those subjects for whom the presence of the line served as the only reliable signal of reinforcement. The presence of the color cue in training, therefore, served to overshadow the line.

C. Cue Neutralization: Overshadowing of Irrelevant by Relevant Stimuli

As in Pavlov's original experiments, the term "overshadowing" refers to the observation that one stimulus, presumably by virtue of its greater intrinsic salience, may detract from the control acquired by another. Pavlov attributed the occurrence of overshadowing to differences in the intensity or salience of overshadowing and overshadowed cues. But there is also evidence that differences in the validity of different stimuli may produce overshadowing (Wagner, 1969a): a given stimulus may acquire relatively litle associative strength if it is reinforced only in conjunction with another, more valid, predictor of reinforcement.

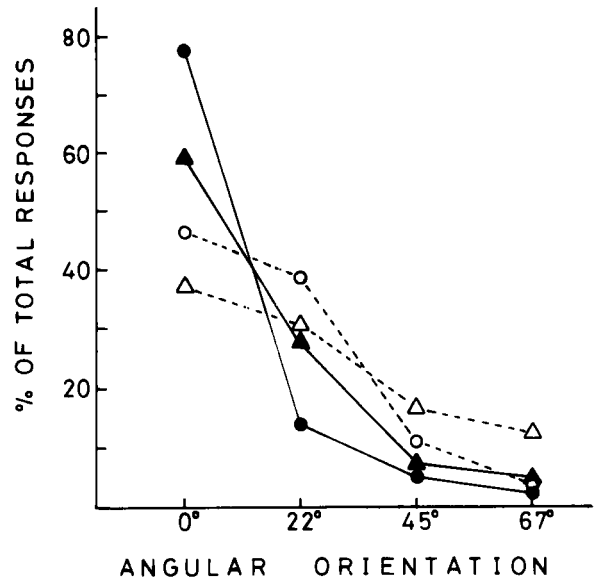

FIGURE 4 Relative generalization gradients to lines of different orientation, after reinforcement for responding to a vertical line (0°): (●) train and test line only; (○) train line, test compound; (▲) train compound, test line only; (△) train and test compound. (After Farthing, 1972.)

1. *Effects of Discrimination Training on Generalization*

This observation provides a simple and convincing account of one of the most ubiquitous results of experiments on generalization. In addition to showing that nondifferential reinforcement in the presence of a 1000-Hz tone might result in little or no apparent control by the tone, Jenkins and Harrison (1960) also reported that discrimination training between the presence and absence of the tone resulted in reliably sloping gradients when the frequency of the tone was varied. This finding, although superficially consistent with Lashley and Wade's (1946) claim that discrimination training was necessary for the appearance of sloping gradients, is clearly inconsistent with their argument that such training must be between different values of the test dimension. The point of Jenkins and Harrison's experiment is that discrimination training along one dimension—between the presence and absence of a tone, increased the slope of a gradient along another dimension—that of auditory frequency.

In order to explain the effect of discrimination training observed by Jenkins and Harrison, it is necessary to recall our earlier analysis of their failure to find significant evidence of auditory control following nondifferential reinforcement. The suggestion was that, even if the tone did acquire control during nondifferential reinforcement, such control would have been

masked by the presence of other, more salient stimuli, such as the key light.

Differential reinforcement should ensure that the tone is a relatively more valid predictor of reinforcement than are other features of the experimental situation, which remain constant on positive and negative trials. Discrimination training therefore should result in some overshadowing of these irrelevant stimuli by the more valid tone. If it is the masking effect of these other stimuli that prevents the tone from controlling responding after nondifferential reinforcement, such overshadowing of potential masking stimuli will ensure the appearance of reliable auditory gradients.

Again, there is relatively direct evidence to support this contention. Miles, Mackintosh, and Westbrook (1970) trained pigeons to peck a key illuminated with light of a given color in the presence of a tone. In the absence of any discriminative training, they found that the color of the key exercised strong control over responding, but variations in the tone had little effect on behavior. They then instituted discrimination training, with reinforced trials to the tone–color compound alternating with nonreinforced trials, on which white noise was substituted for the tone, but the key was illuminated with the same color as on positive trials. Subsequent test trials now indicated that the tone had acquired good control over responding, but that the control exercised by the key light had correspondingly declined.

2. *Irrelevant Stimuli in Simultaneous Discriminations*

One important reason, then, why discrimination training so often sharpens generalization gradients is that differential reinforcement correlated with one set of stimuli results in the suppression of control by stimuli less well correlated with reinforcement. Irrelevant stimuli fail to maintain control of behavior because they are eventually overshadowed by relevant stimuli. The same principle presumably applies to the case of simultaneous discrimination learning. Rats develop position habits when trained on visual discriminations, because spatial stimuli, although less well correlated with reinforcement than are the experimenter's discriminative stimuli, are intrinsically more salient and hence gain initial control. Eventually, however, the greater validity of the relevant visual stimuli enables them to suppress control by spatial stimuli, the rat abandons the position habit, and selects the positive stimulus.

There are doubtless many other irrelevant stimuli in simultaneous visual discriminations, just as there are in successive discriminations. Some of them, such as the shape and size of the two doors in a jumping stand or discrimination box, or the two keys in a pigeon's box, will be common to both S+ and S—. The existence of such stimuli has always posed prob-

lems for a large class of theoretical models: for stimulus-sampling theory, for example, it has not been clear how choice behavior could become perfect when a significant proportion of stimulus elements, which it was necessary to regard as common to S+ and S−, could not be conditioned exclusively to the correct response. The only satisfactory solution is to say that such elements eventually become "adapted" or "neutralized." These, of course, are just descriptive terms. The *process* involved must be one of overshadowing: common elements, as relatively invalid predictors of reinforcement, are overshadowed by those elements correlated with the occurrence of reinforcement.

D. Conclusion

The principle of overshadowing, therefore, originally discovered in experiments on classical conditioning, provides an important key to the solution of some of the more important problems of discrimination learning. That it should have taken so long to recognize this solution is presumably to be attributed to the theoretical problem raised by acceptance of the principle of overshadowing. Explanations of overshadowing are discussed at length both by Rudy and Wagner (Chapter 7, Volume 2) and by Rescorla (Chapter 1, Volume 2), and there is no need to provide further comment here. It is only necessary to point out that traditional theories of conditioning (and, by implication, conditioning-extinction theories of discrimination learning) cannot readily account for the observation that the associative strength of a stimulus depends not just on the salience of that stimulus, and on its correlation with reinforcement or validity, but also on the salience and validity of all other stimuli present at the same time. That overshadowing is inconsistent with traditional conditioning *theory*, however, does not imply that we need one theory of associative learning for simple conditioning, and another for discrimination learning.

V. TRANSFER PROCESSES

We have seen that the phenomena of selective association, so far from being unique to discrimination learning, have in recent years been documented more thoroughly, and analyzed more intensively, in studies of classical conditioning. This does not, however, rule out the possibility that there are other processes, not typically tapped by studies of classical conditioning, which are brought into play only by the more complex requirements of studies of discrimination learning. A not unreasonable suggestion is that even if both discrimination learning and simple conditioning involve similar associations between stimuli and reinforcers, the former involves additional learning mechanisms not involved in the latter. Thus, one might accept

that in both conditioning and discrimination learning stimuli are established as signals for reinforcement in accordance both with their absolute and with their relative salience and validity, but still argue that discrimination learning involves more than just learning to approach and avoid specific stimuli associated with reinforcement and nonreinforcement.

A. General Transfer

Such an argument is supported by evidence of transfer between discriminative problems. If training on one discrimination facilitates the learning of a second, this may imply that part of what the subject learns during the course of training on any discrimination somehow transcends the specific stimuli and responses of that problem and their relationship to reinforcement. The argument requires, of course, that such instances of positive transfer cannot be accounted for by appealing to any simple changes in the associative strengths of relevant and irrelevant stimuli resulting from exposure to the initial discrimination. As an example, a pigeon trained to discriminate between vertical and horizontal lines will probably show marked positive transfer if subsequently trained to discriminate between lines of 30 and 60°. This finding, however, may reflect no more than the generalization of appropriate approach and avoidance tendencies from one pair of lines to a second, similar pair. In order to rule out such specific causes of transfer, experimenters have studied the effects of learning one discrimination on performance with entirely new stimuli, which ostensibly bear no relation to the discriminative stimuli to which the subject has earlier learnt to respond. After being trained on a horizontal–vertical discrimination, for example, pigeons might be required to discriminate between red and green circles. As we shall see, however, even here it will be necessary to treat positive results with some caution, for it may still be possible to explain such apparently general transfer effects without appealing to any processes more complex than those of simple associative conditioning.

1. Positive Transfer from the Neutralization of Irrelevant Stimuli

There is, indeed, evidence that if rats are given sufficiently extensive training on a simultaneous visual discrimination between one pair of stimuli, they may learn a subsequent discrimination between an unrelated pair of stimuli more rapidly than an untrained control group (Bitterman & McConnell, 1954; Mandler, 1966). Similarly, rats and pigeons trained on a free-operant successive discrimination may learn a second successive discrimination significantly more rapidly than control subjects given equivalent exposure to the experimental situation, but without the opportunity

to learn a discrimination (Eck, Noel, & Thomas, 1969; Thomas, Miller, & Svinicki, 1971). Frieman and Goyette (1973) have established that such positive transfer may occur not only when the discriminative stimuli of the two problems are from different modalities, but also when the subjects are required to make different instrumental responses in the two problems. Finally, Thomas, Freeman, Svinicki, Burr, and Lyons (1970) have also confirmed the finding originally reported by Honig (1969) that discrimination training between one pair of stimuli may sharpen the generalization gradient formed round an entirely different stimulus if subjects are subsequently required to respond in its presence.

Thomas (1970) has argued that these experiments provide evidence for the transfer of a state of general attentiveness, rather loosely defined as an increased readiness to attribute changes in reinforcement to changes in any set of stimuli the experimenter chooses to manipulate. The results of these experiments, however, seem analogous to those of experiments on discrimination training and generalization, discussed in the preceding section (IV.C.1), and they may therefore be susceptible to a similar explanation. An observed increase in the probability of control by a new set of relevant stimuli may simply be a consequence of a decrease in the probability of control by irrelevant stimuli common to both initial and subsequent discriminations. Discrimination training, as we have seen, will result in the eventual suppression of control by irrelevant stimuli. Nondifferential reinforcement, by contrast, may ensure that salient irrelevant stimuli acquire strong control over performance. When subjects are shifted to a new problem, the rate at which the new relevant stimuli become established as signals for reinforcement will depend on the extent to which they have to compete with these irrelevant stimuli. Thus, training on one discrimination, by neutralizing such irrelevant stimuli, may facilitate the acquisition of a new discrimination, even if there is no relation between the relevant stimuli of the two problems. Since there is a very wide range of potentially interfering irrelevant stimuli in any experimental situation, ranging from those produced by the apparatus (house lights, masking noises, etc.), through those associated with the performance of the required instrumental response (feedback from pecking, lever pressing, or running), to those arising from the subject's internal state (level of motivation, etc.), it is difficult to prove that any particular case of positive transfer is not due to the neutralization of some set of irrelevant stimuli.

2. *Learning Sets*

Although simple cases of transfer between one problem and another, however, may depend on such an effect, there are other cases of transfer that at first sight seem to pose a greater problem for this analysis. The

obvious example is the formation of learning sets in monkeys (Harlow, 1949). A monkey trained on a series of different visual discriminations, each one for a small number of trials, will eventually learn to solve each problem in a single trial—even though the relevant stimuli of the new problem are quite different from those of earlier problems (Schrier, 1971). But even in this case, Harlow (1959), at least, has argued that the positive transfer displayed is a consequence of a reduction in control by irrelevant stimuli. Harlow suggested that monkeys come to the experimental situation with a variety of erroneous response tendencies or "error factors" (which might include a tendency to respond in terms of position, or to persevere with an initially preferred object), and that the gradual emergence of one-trial solutions to each new problem is due to the gradual elimination of these error factors. Although it is tempting to suppose that the learning-set experienced monkey has achieved a new insight into the experimental situation, which enables him to achieve one-trial solutions, the implication is that the reinforcement contingencies of a single trial would always be sufficient to produce essentially perfect choice of the correct alternative on Trial 2, were their effects not masked or overshadowed by irrelevant aspects of the situation.

Once again it is not easy to show that this type of analysis is inadequate. Harlow himself, it is true, lays more stress on the importance of learning to suppress error factors than on learning which alternative is correct; this emphasis is almost certainly wrong, for excellent learning-set performance has been produced by training regimes which could never have suppressed all error factors (e.g., Schusterman, 1962; Warren, 1965). But with greater emphasis on the positive learning produced by exposure to the contingencies of reinforcement associated with the discriminative stimuli, the theory is a powerful one.

There is, however, one aspect of learning-set data that is not touched on by this analysis. Although there are always problems in comparing the performance of different animals in supposedly comparable situations, there is good reason to suppose that animals differ in their capacity to form learning sets. Primates such as marmosets and squirrel monkeys are notably less proficient than macaque monkeys and apes (Warren, 1965). Although dolphins (Herman & Arbeit, 1973) and corvid birds (Hunter & Kamil, 1971) have shown good learning-set performance, this is true of few other nonprimate species. Pigeons and rats, for example, while entirely proficient at simple discriminative tasks, and, in the case of rats at least, capable of learning the reversal of an old problem in a single trial, have not shown reliable evidence of one-trial learning of *new* problems (Warren, 1965). Thus, an explanation of such learning sets that appeals to simple, general principles of learning, derived largely from research on pigeons and rats, seems somewhat inappropriate.

Without more decisive evidence of the reality of differences in the capacity to form learning sets, and without a clearer understanding of what such differences might mean, it would be rash to make too much of this argument. Nevertheless, it is perhaps unlikely that the explanation of any such differences will be readily apparent from the application of a simple, conditioning-extinction theory of discrimination learning. To this extent, therefore, the argument is an important one, pointing as it does to the possibility that some instances of discrimination learning may require more complex theories for their analysis.

B. Dimensional Transfer

Transfer between discriminations with quite independent relevant stimuli has not received the attention which it deserves. It is still too soon, therefore, to assert with any confidence that we understand all the processes involved. The study of more specific instances of transfer, however, has been pursued with at least intermittent diligence for some years. In the field of animal discrimination learning, the classic experiments are those of Lawrence (1949, 1950), and a large body of subsequent research has been reviewed by Sutherland and Mackintosh (1971). There has also been a considerable amount of research with human subjects, frequently with young children (Shepp & Turrisi, 1966; Wolff, 1967).

1. Paradigms for Studying Dimensional Transfer

Lawrence's experiments on the acquired distinctiveness of cues were intended to discover whether training on one discrimination with a particular set of relevant stimuli would facilitate the learning of a second discrimination, with the same set of relevant stimuli but with different response requirements. He found that a rat trained on a simultaneous black–white discrimination (to approach black and avoid white) would learn a successive conditional black–white discrimination (to go to the left to two black stimuli, and to the right to two white stimuli) more rapidly than a rat trained on a simultaneous discrimination with a different set of relevant stimuli. Lawrence also found that rats would learn the reversal of a simultaneous black–white discrimination more rapidly, the more experience they had previously received with brightness discriminations. This finding was subsequently confirmed by a series of studies of the effects of overtraining on discrimination reversal (Mackintosh, 1965a, 1969; Reid, 1953).

The most popular research strategy in work with children has been to compare the speed of learning reversal and nonreversal shifts, or intra- and extradimensional shifts. The design of such experiments is illustrated in Figure 5. A group of subjects is trained on a color discrimination with

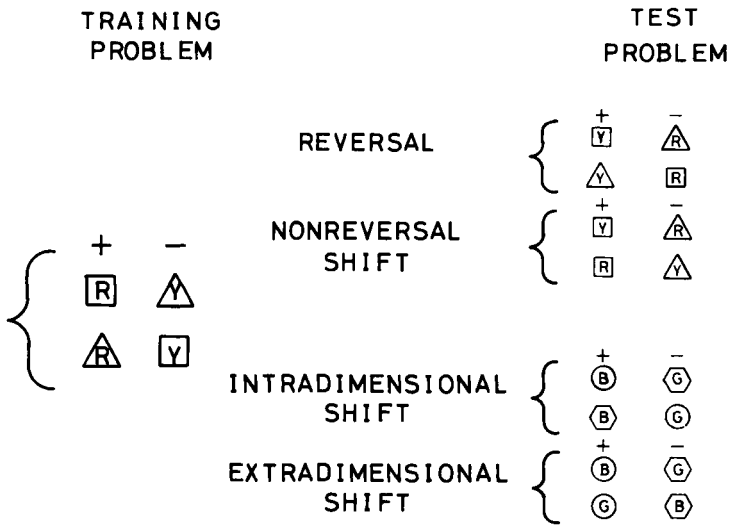

FIGURE 5 Schematic representation of stimuli and reinforcement contingencies for reversal and nonreversal shifts and for intradimensional and extradimensional shifts.

red positive and yellow negative, where the shape of the stimuli (square or triangle) is irrelevant to the correct solution. The reversal of this problem requires subjects to select yellow rather than red, still regardless of shape; a nonreversal shift requires them to select one shape (the square) rather than the other, regardless of color. Children of most ages, but few animals, tend to find the reversal easier than the nonreversal shift. For the comparison of intradimensional and extradimensional shifts, the actual stimuli are changed from the first to the second problem, in the present example from red and yellow squares and triangles to blue and green circles and hexagons. In the intradimensional shift, the dimension relevant in the first problem (color in our example) remains relevant in the second, and subjects are required to select blue rather than green, regardless of shape. In the extradimensional shift, the formerly irrelevant dimension becomes relevant, and subjects are required to select the circle rather than the hexagon, regardless of color. With very few exceptions, experiments with both children and animals (monkeys, rats, pigeons) have found that intradimensional shifts are easier than extradimensional shifts.

2. *Analysis of Dimensional Transfer*

The study of general transfer between unrelated discrimination problems has been hampered by the difficulty of finding an appropriate control group. If subjects trained on one discrimination learn a second faster than

a naive control group, this may be not so much a consequence of having learned any complex hypotheses or discriminative strategies, but rather a case of general habituation to the experimental situation. Although this particular problem may be avoided by using control groups trained in the apparatus on some random schedule of reinforcement and nonreinforcement, it now becomes difficult to decide whether a difference between experimental and control groups is due to the former having learned something of benefit, or to the latter having learned something detrimental to future discriminative performance. These problems do not arise in studies of stimulus-specific or dimensional transfer. In many cases, indeed, the control group in such studies is precisely one that receives training on an unrelated discrimination in the first phase of the experiment. The question at issue, then, is whether subjects whose prior training is on a discrimination where the relevant stimuli are the same as, or from the same dimension as, the relevant stimuli of the test problem, will show superior performance on that test.

Without reviewing this evidence in great detail, it is fair to suggest that this question has in many cases been answered in the affirmative. Rats learn a conditional brightness discrimination faster if they have previously learned a simultaneous brightness discrimination rather than a simultaneous tactile discrimination (e.g., Lawrence, 1949); intradimensional shifts are generally easier than extradimensional shifts (e.g., Schwartz, Schwartz, & Tees, 1971). The most contentious questions are whether overtraining facilitates reversal, and whether reversals are easier than nonreversal shifts. Here there are undoubtedly contradictory findings. Rather than entering into a lengthy discussion of the nature and causes of these discrepancies, however, it is more appropriate here to consider the general implications of the evidence. What is the logic underlying all these experiments, and what theoretical conclusions have they been thought to suggest?

The argument in studies of acquired distinctiveness and dimensional shifts has been that any superiority of experimental to control groups cannot be attributed to the excitatory or inhibitory conditioning to S+ and S− that occurred during the first phase of training, but must reflect a change in the rate at which a specific class of stimuli can be established as signals for action. The most common interpretation of these data is, of course, in terms of theories of selective attention. The assumption underlying such theories is that subjects learn to attend to or ignore particular stimuli in a manner determined by the correlation of those stimuli with reinforcement, and, having learned to attend to one set of relevant stimuli, will learn a subsequent discrimination in which those or similar stimuli are relevant more rapidly than subjects who have not had this prior experience (Sutherland & Mackintosh, 1971).

The crucial step in the argument from data to theory is that the transfer observed cannot be due to the specific response tendencies established as a result of excitatory and inhibitory conditioning in the first phase of the experiment. In the case of dimensional shifts this inference seems reasonable, and has never been seriously disputed; these studies therefore provide the most compelling evidence for this class of theory. Lawrence's experiments on acquired distinctiveness, however, are perhaps more equivocal. Siegel (1967) has argued that the response requirements of simultaneous and successive conditional brightness discriminations may be more similar than has usually been thought. His suggestion was as follows: when solving a simultaneous brightness discrimination, instead of learning to approach black and avoid white regardless of their position, the rat might adopt the habit of always looking right at the choice point, and thus learn to approach black on the right and avoid white on the right.[2] This, however, is closely related to one possible solution of the successive conditional discrimination: here the rat, instead of going right to two black stimuli and left to two white stimuli, might also adopt the habit of initially looking right, and learning to approach black on the right and avoid white on the right. Whether or not Siegel's analysis is applicable to all experiments on acquired distinctiveness (a point disputed by Sutherland & Mackintosh, 1971), the *possibility* of such an analysis illustrates some of the drawbacks inherent in using experiments on instrumental discrimination learning as evidence for dimensional transfer.

Rather similar problems arise in the analysis of reversal learning. The finding that overtraining on a brightness discrimination may paradoxically facilitate the learning of the reversal of the discrimination, seems hard to reconcile with a theory which assumes that discrimination learning involves nothing more than the establishment of excitatory and inhibitory conditioning to S+ and S−. According to theories of selective attention, on the other hand, although overtraining would be expected to increase the discrepancy between the associative strengths of S+ and S−, and thereby retard reversal, it may also increase the probability of attending to the relevant brightness stimuli and thereby facilitate reversal. If the latter effect is greater than the former, the overall effect of overtraining on reversal may be beneficial (Sutherland & Mackintosh, 1971). Although the *general* logic of this explanation seems reasonable enough, it is another matter to prove that the facilitative effect of overtraining is to be identified with the type of attentional process postulated by these theories. It is difficult to rule out the possibility that overtraining somehow alters the way in which

[2] The reader may recall that such a possibility was envisaged earlier (Section III.C.2). Although most accounts of the nature of simultaneous discriminations assume that subjects inspect both alternatives on all trials, we noted that this might not always be true.

subjects orient toward or sample one or other of the discriminative stimuli, thus increasing their ability to react to any change of contingency. Several writers have advanced explanations of this nature (e.g., Hall, 1974b; Mandler, 1966; Siegel, 1967).

The problem with these explanations is that they tend to be specific to particular experimental situations and paradigms. The major strength of theories of attention, it is reasonable to argue, is that they can explain a very wide range of phenomena, without recourse to special ad hoc assumptions. Nevertheless, the available data do not really permit one to dismiss these more specific, rival accounts out of hand. The problem arises from the interaction between the subjects' behavior and their contact with the experimenter's stimulus-reinforcer contingencies that is an inevitable part of any instrumental experiment. The experimenter and the theorist (especially, perhaps, the latter) are inclined to accept a particular, simplified description of the nature of the experimental contingencies, and how they may be assumed to impinge on the subject. But in instrumental experiments, subjects typically control their own exposure to any discriminative stimuli. It would, of course, be idle to pretend that we know that our theoretical description of experiments on classical conditioning coincides with the subjects' own description, but it seems probable that, at least on the stimulus side, the experimenter has better control over the events of a classical experiment than of an instrumental experiment, and therefore better grounds for confidence that his description matches the subjects'. The freedom granted to the subject in a typical instrumental experiment may permit a more complex interaction between behavior, stimulation, and reinforcement than the theorist would wish.

C. Transfer in Classical Conditioning: Latent Inhibition and Learned Irrelevance

It is, then, of some importance to know whether those instrumental results, thought to provide evidence for the transfer of attention from one discriminative situation to another, are readily obtained in analogous classical experiments. Unfortunately, the data are not available. Theories of selective attention have been developed for the data of instrumental discrimination learning, and experiments intended to test such theories have been almost exclusively instrumental. Thus, there are no studies of acquired distinctiveness or dimensional shifts in classical conditioning. The only experiment on overtraining and reversal of which I am aware found no effect, but this negative result loses some of its significance when one considers the relatively narrow range of parameters necessary for the appearance of the effect in instrumental experiments, and the large number of studies in which it

has failed to appear (Sutherland & Mackintosh, 1971). There is, indeed, one feature of most classical experiments that might well militate against the appearance of any of these transfer effects. In terms of theories of selective attention, positive transfer is the result of an increase in the probability of attending to, and hence learning about, a particular set of stimuli, produced by prior experience with those, or similar stimuli. As such, it can be expected only in situations where the stimuli are not so salient that they automatically command a high probability of attention. But the stimuli typically used in classical conditioning, bright lights or loud tones in an otherwise relatively featureless environment, are precisely of this sort. Thus, it is quite possible that no amount of relevant pretraining would serve to increase the probability of attending to a typical CS, and that positive transfer would never be detected.

If this were true, then one might expect studies of classical conditioning to have provided evidence of the converse: if relevant pretraining will not increase the probability of attending to a highly salient CS, perhaps a different form of pretraining will decrease the probability of attention, and hence interfere with subsequent learning. Just such an effect has indeed been observed in a large number of experiments. Lubow and Moore (1959) reported that nonreinforced preexposure to a stimulus would significantly retard conditioning when that stimulus subsequently served as a CS. They termed the effect "latent inhibition," suggesting that this retardation of conditioning was the consequence of the establishment of some response to the preexposed stimulus incompatible with the to-be-established CR. There is good evidence that this explanation cannot be correct, and Lubow himself later rejected it (Lubow, 1973). The critical observation, reported by Rescorla (1971), is that nonreinforced preexposure interferes with the development of both excitatory and inhibitory conditioning. The slow acquisition of a specific CR to a CS may be due to some incompatible response. The slow development of inhibitory conditioning to a CS, however, implies that the CS continues to elicit the CR: the conditioning of an incompatible response to the CS during nonreinforced preexposure would necessarily facilitate rather than retard the development of inhibitory conditioning. The implication, therefore, is that subjects may learn that a stimulus signals no event of consequence during nonreinforced preexposure, and that such learning will interfere with conditioning when the stimulus does signal something—whether it is the presence or the absence of a reinforcer. Such a reduction in the rate of conditioning to a particular stimulus is the hallmark of a reduction in the probability of attending to that stimulus in the sense in which that term has been employed here.

Nonreinforced preexposure is not the only way of reducing attention to a stimulus. There is evidence of a more powerful and specific effect

which may be termed "learned irrelevance": exposure to uncorrelated presentations of a conditioned stimulus and unconditioned stimulus very markedly interferes with subsequent conditioning when the two are correlated (Gamzu & Williams, 1971; Mackintosh, 1973). Thus, we can safely conclude that the rate at which a CS can be established as a signal for a UCS is not a fixed, static consequence of the salience or intensity of that CS, or of its current correlation with reinforcement, but can be significantly affected by the subject's past experience.

With very few exceptions, latent inhibition and learned irrelevance have been studied in classical rather than in instrumental experiments. Halgren (1974) has confirmed Rescorla'a results by showing that nonreinforced prexposure to a stimulus will retard the acquisition of a free-operant successive discrimination in which that stimulus serves either as S+ or as S−; while Mellgren and Ost (1969) have established that exposure to uncorrelated presentations of a stimulus and reinforcer will aslo significantly retard subsequent free-operant discrimination learning. Very few experiments on discrete-trial discrimination learning have reported such results (Bateson & Chantrey, 1972). Several studies, indeed, have suggested that in this situation preexposure may actually facilitate discrimination learning (Bateson, 1973; Gibson & Walk, 1956).

Whatever the explanation of this apparent discrepancy, it serves to remind us of the gulf that separates these two areas of research. Moreover, it may help to explain why the phenomena of latent inhibition and learned irrelevance, although suggesting that attention changes with experience, are not readily explained by theories of selective attention developed with the data of instrumental learning in mind. According to such theories, the probability of attending to stimuli correlated with reinforcement will increase over a series of trials, and may decrease again if those stimuli are no longer correlated with reinforcement. But none of these theories contains machinery for decreasing the probability of attention below an initial base-line level, when a stimulus is presented in the absence of correlated changes in reinforcement. Although studies of acquired distinctiveness and dimensional transfer and experiments on latent inhibition and learned irrelevance have employed entirely different paradigms and procedures, it is hard to believe that such symmetrical effects do not reflect the operation of a single, fundamental mechanism. That mechanism must explain how it is that the establishment of a stimulus as a signal for reinforcement is affected by the subject's past experience with that stimulus, being facilitated by experience of a correlation between stimulus and reinforcer, and interfered with by experience of no correlation between the two. That there is no theory with the requisite properties is a sad reflection on the self-imposed restrictions under which experimental psychologists continue to operate.

VI. CONCLUSION

It is perhaps unnecessary to reiterate at length the central argument that has, in intention at any rate, underlain the diverse issues discussed in this chapter. My main point has been a simple one: if experiments on discrimination learning may be defined as those in which different schedules of reinforcement are correlated with different stimulus situations, then discrimination learning can be reduced to the association of those stimuli with their respective schedules. Since typical experiments correlate one stimulus with reinforcement and a second with the omission of reinforcement, conditioning-extinction theory provides the most plausible account of discrimination learning. The problem with traditional conditioning-extinction theory is that it has taken too simple a view of the processes of conditioning and extinction. Both in the simplest cases of conditioning, and in the most complex cases of discrimination learning, there is evidence that stimuli in some sense compete for association with reinforcement, and that subjects may learn to attend to and ignore stimuli in accordance with their correlation with reinforcement.

My main thesis, therefore, has been that there may be few processes unique to discrimination learning that cannot also be studied in what are traditionally regarded as simpler situations. This does not, of course, imply that we should study only those simpler situations, for the generality of concepts across different situations can only be established by comparison of those situations. It does imply that theories of discrimination learning should be related to other theories—perhaps, indeed, that there should be no such entity as a theory of discrimination learning as such. If it also seems to imply that there are few principles specifically derived from the study of discrimination learning that may illuminate yet more complicated problems, this is only because we may hope that a richer theory of "simple" associative learning will itself contain sufficient insights and complexities to illuminate those areas. There will then be a real chance of breaking down some of the conceptual barriers that have for too long compartmentalized the study of associative learning.

ACKNOWLEDGMENTS

The writing of this chapter was supported by a grant from the U.K. Science Research Council. I am grateful to A. Dickinson, G. Hall, E. M. Macphail, D. L. Medin, and in particular to W. K. Estes for their comments on an earlier draft.

REFERENCES

Bateson, P. P. G. Internal influences on early learning in birds. In R. A. Hinde & J. S. Hinde (Eds.), *Constraints on learning*. London and New York: Academic Press, 1973. Pp. 101–116.

Bateson, P. P. G., & Chantrey, D. F. Retardation of discrimination learning in monkeys and chicks previously exposed to both stimuli. *Nature*, 1972, **237**, 173–174.

Bitterman, M. E., & McConnell, J. V. The role of set in successive discrimination. *American Journal of Psychology*, 1954, **67**, 129–132.

Black, R. W. Shifts in magnitude of reward and contrast effects in instrumental and selective learning. *Psychological Review*, 1968, **75**, 114–126.

Boakes, R. A., Halliday, M. S., & Poli, M. Response additivity: Effects of superimposed free reinforcement on a variable-interval baseline. *Journal of the Experimental Analysis of Behavior*, 1975, **23**, 177–191.

Brown, P. L., & Jenkins, H. M. Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, 1968, **11**, 1–8.

Eck, K. O., Noel, R. C., & Thomas, D. R. Discrimination learning as a function of prior discrimination and nondifferential training. *Journal of Experimental Psychology*, 1969, **82**, 156–162.

Farthing, G. W. Overshadowing in the discrimination of successive compound stimuli. *Psychonomic Science*, 1972, **28**, 29–32.

Freeman, F., & Thomas, D. R. Attention vs cue utilization in generalization testing. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, 1967.

Frieman, J., & Goyette, C. H. Transfer of training across stimulus modality and response class. *Journal of Experimental Psychology*, 1973, **97**, 235–241.

Gamzu, E., & Williams, D. R. Classical conditioning of a complex skeletal response. *Science*, 1971, **171**, 923–925.

Gibson, E. J., & Walk, R. D. The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *Journal of Comparative and Physiological Psychology*, 1956, **59**, 239–242.

Gray, J. A. Stimulus intensity dynamism. *Psychological Bulletin*, 1965, **63**, 180–196.

Grice, G. R. Stimulus intensity and response evocation. *Psychological Review*, 1968, **75**, 359–373.

Gulliksen, H., & Wolfle, H. L. A thoery of learning and transfer: I. *Psychometrika*, 1938, **3**, 127–149.

Halgren, C. R. Latent inhibition in rats: Associative or nonassociative? *Journal of Comparative and Physiological Psychology*, 1974, **86**, 74–78.

Hall, G. Strategies of simultaneous discrimination learning in the pigeon. *Quarterly Journal of Experimental Psychology*, 1974, **26**, 520–529. (a)

Hall, G. Transfer effects produced by overtraining in the rat. *Journal of Comparative and Physiological Psychology*, 1974, **87**, 938–944. (b)

Hanson, H. M. Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, 1959, **58**, 321–334.

Harlow, H. F. The formation of learning sets. *Psychological Review*, 1949, **56**, 51–65.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 2. New York: McGraw-Hill, 1959. Pp. 492–537.

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

Heinemann, E. G., & Rudolph, R. L. The effect of discriminative training on the gradient of stimulus generalization. *American Journal of Psychology*, 1963, **76**, 653–658.

Hemmes, N. S. Behavioral contrast in pigeons depends upon the operant. *Journal of Comparative and Physiological Psychology*, 1973, **85**, 171–178.

Herman, L. M., & Arbeit, W. R. Stimulus control and auditory discrimination learning sets in the bottlenose dolphin. *Journal of the Experimental Analysis of Behavior*, 1973, **19**, 379–394.

Honig, W. K. Prediction of preference, transposition, and transposition-reversal from the generalization gradient. *Journal of Experimental Psychology,* 1962, **64,** 239–248.

Honig, W. K. Attentional factors governing the slope of the generalization gradient. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning.* London and New York: Academic Press, 1969. Pp. 35–62.

Hull, C. L. Stimulus intensity dynamism (V) and stimulus generalization. *Psychological Review,* 1949, **56,** 67–76.

Hull, C. L. *A behavior system.* New Haven, Connecticut: Yale University Press, 1952.

Hunter, M. W., III, & Kamil, A. C. Object discrimination learning set and hypothesis behavior in the northern bluejay (*Cyanocitta cristata*). *Psychonomic Science,* 1971, **22,** 271–273.

Jenkins, H. M. Noticing and responding in a discrimination based on a distinguishing element. *Learning and Motivation,* 1973, **4,** 115–137.

Jenkins, H. M., & Harrison, R. H. Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology,* 1960, **59,** 246–253.

Kamin, L. J. Temporal and intensity characteristics of the conditioned stimulus. In W. F. Prokasy (Ed.), *Classical conditioning: A symposium.* New York: Appleton-Century-Crofts, 1965. Pp. 118–147.

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. Pp. 9–33.

Kamin, L. J. Predictability, surprise, attention and conditioning. In B. Campbell & R. Church (Eds.), *Punishment and aversive behavior.* New York: Appleton-Century-Crofts, 1969. Pp. 279–296.

Köhler, W. Nachweis einfacher Strukturfunktionen beim Schimpansen und beim Haushuhn: Über eine neue Methode zur Untersuchung des bunten Farbensystems. *Abhandlungen der Preussischen Akademie der Wissenschaften,* 1918, 2, 1–101. Translated as: Simple structural functions in the chimpanzee and in the chicken. In W. D. Ellis (Ed.), *A source book of Gestalt psychology.* New York: Harcourt, Brace, 1939. Pp. 217–227.

Lashley, K. S. An examination of the continuity theory as applied to discriminative learning. *Journal of General Psychology,* 1942, **26,** 242–265.

Lashley, K. S., & Wade, M. The Pavlovian theory of generalization. *Psychological Review,* 1946, **53,** 72–87.

Lawrence, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology,* 1949, **39,** 770–784.

Lawrence, D. H. Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology,* 1950, **40,** 175–188.

Logan, F. A. A note on stimulus intensity dynamism (V). *Psychological Review,* 1954, **61,** 77–80.

Lovejoy, E. *Attention in discrimination learning.* San Francisco: Holden-Day, 1968.

Lubow, R. E. Latent inhibition. *Psychological Bulletin,* 1973, **79,** 398–407.

Lubow, R. E., & Moore, A. U. Latent inhibition: The effect of nonreinforced preexposure to the conditioned stimulus. *Journal of Comparative and Physiological Psychology,* 1959, **52,** 415–419.

Mackintosh, N. J. Selective attention in animal discrimination learning. *Psychological Bulletin,* 1965, **64,** 124–150. (a)

Mackintosh, N. J. Incidental cue learning in rats. *Quarterly Journal of Experimental Psychology,* 1965, **17,** 292–300. (b)

Mackintosh, N. J. Further analysis of the overtraining reversal effect. *Journal of Comparative and Physiological Psychology, Monograph*, 1969, **67**, No. 2, Part 2.

Mackintosh, N. J. Stimulus selection: learning to ignore stimuli that predict no change in reinforcement. In R. A. Hinde & J. S. Hinde (Eds.), *Constraints on learning*. London and New York: Academic Press, 1973. Pp. 75–96.

Mackintosh, N. J., & Honig, W. K. Blocking and attentional enhancement in pigeons. *Journal of Comparative and Physiological Psychology*, 1970, **73**, 78–85.

Mahut, H. The effect of stimulus position on visual discrimination by the rat. *Canadian Journal of Psychology*, 1954, **8**, 130–138.

Mandler, J. M. Behavior changes during overtraining and their effects on reversal and transfer. *Psychonomic Monograph Supplements*, 1966, **1**, 187–202.

Medin, D. L. A context theory of discrimination learning. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, New York: Academic Press, in press.

Mellgren, R. L., & Ost, J. W. P. Transfer of Pavlovian differential conditioning to an operant discrimination. *Journal of Comparative and Physiological Psychology*, 1969, **67**, 390–394.

Miles, C. G., Mackintosh, N. J., & Westbrook, R. F. Redistributing control between the elements of a compound stimulus. *Quarterly Journal of Experimental Psychology*, 1970, **22**, 478–483.

Moore, J. W. Stimulus control: Studies of auditory generalization in rabbits. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts, 1972. Pp. 206–230.

Newman, F. L., & Baron, M. R. Stimulus generalization along the dimension of angularity. *Journal of Comparative and Physiological Psychology*, 1965, **60**, 59–63.

Pavlov, I. P. *Conditioned reflexes*. London and New York: Oxford University Press, 1927.

Pear, J. J., & Wilkie, D. M. Contrast and induction in rats on multiple schedules. *Journal of the Experimental Analysis of Behavior*, 1971, **15**, 289–296.

Perkins, C. C., Jr. The relation between conditioned stimulus intensity and response strength. *Journal of Experimental Psychology*, 1953, **46**, 225–231.

Reid, L. S. The development of noncontinuity behavior through continuity learning. *Journal of Experimental Psychology*, 1953, **46**, 107–112.

Rescorla, R. A. Summation and retardation tests of latent inhibition. *Journal of Comparative and Physiological Psychology*, 1971, **75**, 77–81.

Rescorla, R. A., & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts, 1972. Pp. 64–99.

Reynolds, G. S. Behavioral contrast. *Journal of the Experimental Analysis of Behavior*, 1961, 4, 57–71.

Riley, D. A. *Discrimination learning*. Boston: Allyn & Bacon, 1968.

Riley, D. A., Goggin, J. P., & Wright, D. C. Training level and cue separation as determiners of transposition and retention in rats. *Journal of Comparative and Physiological Psychology*, 1963, **56**, 1044–1049.

Riley, D. A., Ring, K., & Thomas, J. The effect of stimulus comparison on discrimination learning and transporition. *Journal of Comparative and Physiological Psychology*, 1960, **53**, 415–421.

Rudolph, R. L. Transposition and the post-discrimination gradient: Absolute and relational learning during hue discrimination. Paper presented at the meeting of the Eastern Psychological Association, Boston, 1967.

Rudolph, R. L., & Van Houten, R. Auditory stimulus control in pigeons: Jenkins and Harrison (1960) revisited. *Journal of the Experimental Analysis of Behavior,* in press.

Sachs, L. B. Effects of stimulus comparison during discrimination training on subsequent transposition and generalization gradients. *Psychonomic Science,* 1969, **14,** 247–248.

Schrier, A. M. Extradimensional transfer of learning-set formation in stumptailed monkeys. *Learning and Motivation,* 1971, **2,** 173–181.

Schusterman, R. J. Transfer effects of successive discrimination reversal training in chimpanzees. *Science,* 1962, **137,** 422–423.

Schwartz, R. M., Schwartz, M., & Tees, R. C. Optional intradimensional and extradimensional shifts in the rat. *Journal of Comparative and Physiological Psychology,* 1971, **77,** 470–475.

Shepp, B. E., & Turrisi, F. D. Learning and transfer of mediating responses in discriminative learning. In N. R. Ellis (Ed.), *International review of research in mental retardation.* Vol. 2. New York: Academic Press, 1966. Pp. 85–121.

Siegel, S. Overtraining and transfer processes. *Journal of Comparative and Physiological Psychology,* 1967, **64,** 471–477.

Spence, K. W. The nature of discrimination learning in animals. *Psychological Review,* 1936, **43,** 427–449.

Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychological Review,* 1937, **44,** 430–444.

Spence, K. W. The nature of the response in discrimination learning. *Psychological Review,* 1952, **59,** 89–93.

Spiker, C. C. An extension of Hull-Spence discrimination learning theory. *Psychological Review,* 1970, **77,** 496–515.

Stollnitz, F. Spatial variables, observing responses, and discrimination learning set. *Psychological Review,* 1965, **72,** 247–261.

Sutherland, N. S. The learning of discrimination by animals. *Endeavour,* 1964, **23,** 69–78.

Sutherland, N. S., & Mackintosh, N. J. *Mechanisms of animal discrimination learning.* New York: Academic Press, 1971.

Terrace, H. S. Stimulus control. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application.* New York: Appleton-Century-Crofts, 1966. Pp. 271–344.

Thomas, D. R. Stimulus selection, attention, and related matters. In J. H. Reynierse (Ed.), *Current issues in animal learning.* Lincoln: University of Nebraska Press, 1970. Pp. 311–356.

Thomas, D. R., Freeman, F., Svinicki, J. G., Burr, D. E. S., & Lyons, J. Effects of extradimensional training on stimulus generalization. *Journal of Experimental Psychology, Monograph,* 1970, **83,** No. 1, Part 2.

Thomas, D. R., Miller, J. T., & Svinicki, J. G. Nonspecific transfer effects of discrimination training in the rat. *Journal of Comparative and Physiological Psychology,* 1971, **74,** 96–101.

Wagner, A. R. Stimulus validity and stimulus selection in associative learning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning.* Halifax: Dalhousie University Press, 1969. Pp. 90–122. (a)

Wagner, A. R. Incidental stimuli and discrimination learning. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning.* London and New York. Academic Press, 1969. Pp. 83–111. (b)

Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. Stimulus selection in

animal discrimination learning. *Journal of Experimental Psychology*, 1968, **76,** 171–180.

Warren, J. M. Primate learning in comparative perspective. In A. M. Schrier, H. F. Harlow, & F. Stollnitz (Eds.), *Behavior of nonhuman primates.* Vol. 1. New York: Academic Press, 1965. Pp. 249–281.

Westbrook, R. F. Failure to obtain positive contrast when pigeons press a bar. *Journal of the Experimental Analysis of Behavior*, 1973, **20,** 499–510.

Wolff, J. L. Concept-shift and discrimination-reversal learning in humans. *Psychological Bulletin*, 1967, **68,** 369–408.

Zeaman, D., & House, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency: Psychological theory and research.* New York: McGraw-Hill, 1963. Pp. 159–223.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

6

From Discrimination Learning to Cognitive Development: A Neobehavioristic Odyssey

Howard H. Kendler
Tracy S. Kendler[1]

University of California, Santa Barbara

Kurt Koffka served as a midwife in the birth of neobehaviorism. While listening to Koffka's lecture on Gestalt psychology, Clark L. Hull (1884–1952) was encouraged to conceptualize a new kind of behaviorism:

> At least half of his (Koffka's) time was spent in attacking Watson. . . . I came to the conclusion not that the Gestalt view was sound but rather that Watson had not made out as clear a case for behaviorism as the facts warranted. Instead of converting me to *Gestalt theory*, the result was a belated conversion to a kind of neo-behaviorism—a behaviorism mainly concerned with the determination of the quantitative laws of behavior and their deductive systemization [Hull, 1952a, p. 154].

Hull spent the major part of his career formulating a systematic theory of behavior with testable implications. In his effort he was assisted by numerous disciples, most notably Kenneth W. Spence (1907–1967) and Neal E. Miller (1909–), both of whom operated within the broad methodological and theoretical traditions Hull had established and each in his own way contributed to the development, refinement, and expansion of the neobehavioristic system.

[1] Neither author fully agrees with all points of view expressed in this presentation but each will defend the right of the coauthor to state them.

Neobehaviorism achieved its peak of popularity in the 1940s and early 1950s. In the ten years following the publication of Hull's major work, *Principles of Behavior* (1943), 70% of all the articles published in the *Journal of Experimental Psychology* and the *Journal of Comparative and Physiological Psychology,* two of the most prestigious periodicals of that era, made reference to the writings of Clark Hull (Spence, 1943). But with passing time, and the deaths of both Hull and Spence, neobehaviorism lost some of its distinctiveness and most of its zeal, and the question has been raised as to whether it will persist as an independent force in psychology (Kendler & Spence, 1971).

I. THE MEANING OF NEOBEHAVIORISM

Hull's description of the goal of neobehaviorism—a theoretical integration of quantitative laws of behavior—is not sufficiently comprehensive to characterize neobehaviorism as a distinctive tradition. Although adequate to differentiate neobehaviorism from systematic positions that either eschew deductive theories (Skinner, 1950) or encourage subjectivity (e.g., Giorgi, 1970), Hull's description falls far short of setting neobehaviorism apart from other orientations that aspire to general quantitative formulations of behavior (e.g., Estes, 1959). To appreciate fully the character of neobehaviorism, one must identify specific methodological assumptions and strategic decisions that governed its development. Systematic orientations, such as neobehaviorism, are characterized by a network of commitments (not necessarily logically intertwined), some of which are carefully inspected and deliberately adopted, while others are accepted with little or no critical examination. Only from an historical perspective does it become possible to perceive the methodological and strategic underpinnings of neobehaviorism.

A. Methodological Assumptions of Neobehaviorism

Neobehaviorism's conception of the science of psychology was shaped by the implications of two of its fundamental epistemological assumptions: methodological behaviorism and a deductive model of explanation.

1. Methodological Behaviorism

The uncritical use of the term *behaviorism* has done much to confuse discussions of methodological issues in contemporary psychology. More often than not behaviorism is conceived as a limited and rigid set of methodological prescriptions and theoretical assumptions that, at best, reflects an incomplete picture and, at worst, a gross distortion. Such conceptions ignore both the historical context in which behaviorism emerged (Watson,

1913) and later subvarieties of behaviorism with methodologically opposed positions. For example, the radical behaviorism of Skinner (1969) is strikingly different from the purposive behaviorism of Tolman (1932). The latter position, it needs be remembered, anticipated the cognitive revolution in psychology.

In its most limited sense, behaviorism can be thought of as a *methodological behaviorism,* a passionate devotion to the ideals of scientific objectivity; psychological research, data, and theorizing should be open to public examination. Admittedly the concept of *objectivity* is ambiguous and its role in science controversial (Brush, 1974). Nevertheless, the position adopted here is that objective standards in data collecting and in evaluating theoretical statements have meaning, especially in psychology, the history of which is dotted with recurring controversies about the appropriate role of subjective experience and intuitive convictions. The meaningfulness of objectivity is particularly clear when one compares methodological behaviorism with orientations that espouse the positions that mental states are the primary dependent variable of psychology and/or the investigation of mental events requires abandonment of public criteria for concepts, theories, and explanation.

Of basic importance for the proper understanding of methodological behaviorism is its position on *conscious experience, raw feels, mental states,* or any of the other numerous terms that refer to *subjective experience.* In actual fact, methodological behaviorism, within limits, maintains an open-minded view of the mind. By demanding intersubjective agreement for psychological data, methodological behaviorism rejects as inadequate evidence a person's private experience simply because it is unavailable to the inspection of other observers. Such a methodological stance in no way denies the existence of mental states nor the acceptability of introspective *reports* of mental states as data since such reports are obviously available to public observation. How important introspective reports are is the point at issue among methodological behaviorists. Three general reactions can be identified: introspective reports (1) are better ignored, (2) can be useful in formulating hypotheses, and (3) may reflect causal events.

The negative reaction to verbal reports of phenomenal experience is easily understood within the context of history. Introspection, the fundamental method of structuralism, which was the dominant orientation in psychology when it achieved the status of an independent science, failed to achieve consensual agreement about the facts of psychology. The behavioristic revolution should be viewed primarily as a reaction to structuralism's failures. But why did structuralism fail? Did it fail because conscious experience was considered as a legitimate phenomenon to investigate *or* because of the inadequacies of their particular introspective method? Watson failed to distinguish between these alternatives, and as a result the problem

of conscious experience and methods used to investigate it were *jointly* rejected by the early behaviorists. Later, however, it became apparent that verbal reports are as behavioral as bar presses and therefore are legitimate data. Consequently, evaluations of verbal reports shifted from questions of methodological propriety to those of theoretical significance. It became apparent to some psychologists that verbal reports could be useful. For example, a methodological behaviorist interested in memory processes could not avoid becoming impressed with the subjects' reports of their encoding techniques. Such evidence encouraged the investigation of encoding processes and their incorporation into models of memory. Still another possible tack for a methodological behaviorist was to assign special theoretical import to inferred phenomenal experience because of its significant influence on subsequent behavior, as a person's perception even when false, can play a more important role in behavior than the physically "real" situation. A false heart rate feedback, for example, was more influential in determining a subject's reaction to erotic photographs than his actual heart rate (Valins, 1966).

The purpose of this analysis of verbal reports of phenomenal experience was not to offer recommendations about how such verbal reports are to be used, if at all, but instead to identify their epistemological status within methodological behaviorism. The point at issue is not whether verbal reports are used but instead how they are used. They meet the demands of methodological behaviorism when used as publicly observable responses and/or data from which to infer theoretical processes.

2. *A Deductive Model of Explanation*

The neobehavioristic position with regard to explanation is clear; scientific understanding is achieved when a phenomenon (for example, ontogenetic changes in discrimination-shift behavior, greater resistance to extinction of intermittent relative to continuous reinforcements) can be deduced, that is predicted, either in an a priori or a posteriori fashion, from a set of theoretical propositions. The deductive process is analogous to mathematical proof although its precision can vary from mathematical verification to the logical use of ordinary language. The early neobehaviorists were particularly influential in espousing a deductive model of explanation by cleverly demonstrating how a variety of behavioral phenomena could be deduced from principles of conditioning, for example, maze behavior (Hull 1952b), discrimination learning (Spence, 1936, 1937a), and conflict behavior (Miller, 1944).

The virtues of deductive explanations appear so apparent to theoretically oriented experimental psychologists that it becomes incomprehensible that

any other criterion of understanding could be adopted. A survey of psychology would demonstrate, however, that other criteria of understanding are currently in use: (1) behavioral control, (2) interpretive consistency, and (3) intuitive validity. The various conceptions of understanding exert profoundly different influences on scientific efforts.

An operant conditioning approach adopts behavioral control as its goal while simultaneously eschewing abstract theories. Although the search for deductive explanation and behavioral control generates some overlapping interests and actions (H. H. Kendler, 1970; Scriven, 1956), they nevertheless encourage divergent research practices (D'Amato, 1970).

Understanding achieved through interpretive consistency is illustrated by psychoanalytic explanations of behavior. The behavior of individuals become comprehensible when their actions dovetail with a preconceived pattern, for example, a homosexual panic is conceived as an outgrowth of an unresolved Oedipus complex (H. H. Kendler, 1947). In principle, such pattern interpretations can have deductive capacities. In practice, however, they prove to be unfalsifiable because they "predict" everything and anything (Nagel, 1959). Witness the ease with which different psychoanalytic systems (Freudian, Jungian, Adlerian, etc.), based on conflicting assumptions, offer apparently consistent interpretations of the same behavioral phenomena.

Intuitive validity, the phenomenal conviction that a proposition is true, is a criterion of understanding that seems to be so far removed from the tough-minded discipline of experimental psychology, that its mention appears inappropriate. Nevertheless, the question can be raised whether important assumptions with empirical implications have been made in the past primarily on the basis of their phenomenal appeal instead of their deductive implications, for example, conditioning is the most elementary form of learning, fundamental principles of perception can be phenomenally observed, human language is innate.

Our analysis of different forms of understanding was not offered to encourage controversy in hope of identifying the "true" meaning of scientific understanding. Such a goal cannot be attained for the simple reason that the choice of the criteria for explanation is a matter of preference, not logic. Neither was its purpose to argue that the four kinds of understanding are mutually exclusive; for example, a satisfactory deductive theory obviously provides an interpretive pattern and generates, at least for some, a sense of intuitive understanding. In addition, a deductive theory is frequently the first step in the direction of successful control. Our major aim was to highlight a methodological commitment to a deductive model of explanation which distinguishes neobehaviorism from some other current psychological orientations.

B. The Strategy of Neobehaviorism

Commitments to methodological behaviorism and a deductive model of explanation specify standards and goals of scientific inquiry; strategies represent the techniques used to achieve such objectives. Three strategies that have been most influential in shaping the development of neobehaviorism, and in setting it apart from other approaches are (1) the choice of conditioning phenomena as a source of principles of behavior, (2) the specification of boundary conditions for theoretical formulations, and (3) the use of a stimulus–response conception of behavior.

1. Conditioning as a Source for Principles of Behavior

It should not be too difficult to imagine how Hull, committed to mehodological behaviorism and striving to formulate a deductive theory, would hit upon the apparently reasonable strategy of abstracting principles of behavior from conditioning phenomena. In addition to meeting the standards of objectivity, conditioning appeared to be the simplest form of learning, with classical conditioning illustrating how associations are formed and both classical and instrumental conditioning revealing how they are strengthened. In retrospect, one can see that Hull failed to appreciate the complexity of these assumptions (H. H. Kendler & Spence, 1971). One reason is that he did not distinguish between two possible meanings of simplicity: devoid of complexity and theoretically fundamental. Conditioning initially appeared to many experimental psychologists to be the uncomplicated consequence of a few variables; classical conditioning resulted from the appropriate scheduling of conditioned and unconditioned stimuli, whereas instrumental conditioning was a matter of rewarding desired responses. Being so uncomplicated, it was automatically assumed that conditioning must be fundamental, capable of reflecting basic principles of learning. Although the idea was appealing, it turned out to be unjustified because it was based on the false logic that lack of complexity is equivalent to being theoretically fundamental. Although causally simple, conditioning could nevertheless be, as Tolman suggested (1932), an artificial phenomenon, unrepresentative of other forms of learning. The significant point is that the fundamental nature of conditioning could not be prejudged. Many neobehaviorists did not fully appreciate this point and therefore, with blind faith, pursued the investigation of conditioning with the expectation that such efforts would yield general principles of behavior. Success was not ensured by such a strategy, as witness the failure to discover basic temporal principles of associative learning by searching for the optimal time interval between the conditioned and unconditioned stimuli (Gormezano & Moore,

1969). In defense of this early neobehavioristic strategy, one could argue, that its limitations could never be discovered without making the effort. In any case, the historical verdict for the assumption that a comprehensive theory of learning, much less a theory of behavior, could be based upon conditioning principles, is not proven. This does not necessarily mean that conditioning principles are irrelevant; they could be insufficient.

2. *Boundary Conditions of Theories*

After suggesting that one tactic of neobehavioristic strategy was to assume that principles of behavior could be gleaned from conditioning, it will now appear inconsistent to argue that the specification of theoretical boundary conditions was also characteristic of neobehaviorism. Major orientations in psychology have not always spoken with a single voice, and in the case of neobehaviorism, differing positions were adopted by Hull and Spence. The former, in his preface of *Principles of Behavior* (1943) which mainly reported evidence from conditioning studies, wrote:

> As suggested by the title, this book attempts to present in an objective, systematic manner the primary, or fundamental molar principles of behavior. It has been written on the assumption that all behavior, individual and social, moral and immoral, normal and psychopathic, is generated from the same primary laws: that the differences in the objective behavioral manifestations are due to the differing conditions under which habits are set up and function. Consequently the present work may be regarded as a general introduction to the theory of all behavioral (social) sciences [p. v].

Contrasted with this optimistic view was Spence's caution in suggesting the boundary conditions of his discrimination learning theory. "It is possible only with the advent of verbal processes that the simple mechanism of learning discrimination problems we have proposed, is transcended" (Spence, 1937b, p. 99).

His example encouraged his students (e.g., H. H. Kendler & T. S. Kendler, 1962; Kuenne, 1946), to not only identify the boundary conditions but also to examine the similarities and differences between the theoretical principles used to explain animal learning and those required to explain human cognitive behavior. Spence (1966) attempted to do this himself later in his career in his analysis of cognitive and drive factors in human eyelid conditioning.

3. *Stimulus–Response Conception in Behavior*

A controversial, strategic decision of neobehaviorism was the adoption of a stimulus–response (S–R) conception of behavior. The implications of such a tactic have been argued at great length, frequently generating more

heat than light. The choice of an S–R conception is not difficult to understand, considering historical pressures operating at the time of its adoption. A stimulus–response language, with its emphasis on physicalistic measures of environmental and behavioral events, appeared to be a safeguard against any subjectivism that might infiltrate psychological formulations. In addition, it has been suggested that "S–R language clarifies the experimental and theoretical task of the psychologist by focussing attention on the empirical relationships to be investigated and explained" (H. H. Kendler & Spence, 1971, p. 32). An opposing position would be that S–R language is intrinsically misleading because it limits behavioral analysis to stimulus, response, and associative events, a restriction that prevents an adequate description of behavior, particularly of the kind characterized as "cognitive."

An appropriate evaluation of stimulus–response language must await an historical verdict, and then it is likely that the judgment will neither be simple nor unqualified. Of particular relevance to this presentation are the answers to two questions:

1. Was a simple S–R model, involving only concepts of stimulus, response, and association, appropriate for some forms of behavior?
2. Was it inconsistent, in principle, for a S–R formulation to be expanded to include concepts other than the aforementioned three?

4. Final Comments about the Strategy of Neobehaviorism

Definitive evaluation of strategic decisions are difficult to arrive at for two reasons: (1) the consequences of a particular strategy are not isolated from the effects of other strategies; (2) when a particular strategy encounters difficulty, or fails, one cannot be sure that the strategy itself was at fault or the manner in which it was implemented.

If treated neither with scorn nor reverence, these strategic decisions of neobehaviorism appear to have had both positive and negative consequences. Accepting conditioning as a source of behavioral principles had the virtue of encouraging systematic theorizing to account for the facts of conditioning, while simultaneously creating the myth that, in the face of a good theory of conditioning, the obstacles to understanding behavior would collapse. Spence's caution in suggesting boundary conditions for his animal discrimination-learning theory, although an important step in the right direction, was not articulated with sufficient force or clarity to emphasize or identify fundamental comparative and developmental differences with which theories of learning and behavior were required to cope. Although a stimulus–response conception of behavior had salutary effects on

objective experimental and theoretical practices, it nevertheless provided an oversimplified view of behavior. In essence, the strategy of neobehaviorism opted on the side of parsimony in assuming that errors of oversimplification would be less corrupting and more easily corrected than those of overcomplication.

II. DISCRIMINATION LEARNING: THE PAST

In a conventional discrimination-learning experiment, the subject is presented with two different discriminanda. Responses to one discriminandum are rewarded; responses to the other are either unrewarded or punished. For example, a discrimination-learning experiment in which rats serve as subjects might involve a choice between a black and white alley. If the subject runs down the black alley, he finds food, whereas none is available in the white alley. The subject is said to have learned the discrimination when he consistently chooses the black alley, given that its right or left position is randomized. This method was used originally to measure the sensory capacities of animals; if the animal learns the discrimination, it must be able to discriminate between the stimuli.

A. The Continuity–Noncontinuity Controversy

The focus of interest in discrimination learning shifted for learning theorists from the discrimination process to the learning process, leading ultimately to the continuity–noncontinuity controversy. Noncontinuity theory, originally proposed by Lashley (1929) and Krechevsky (1932), conceived of discrimination learning as a problem-solving process that consisted of successively testing alternative hypotheses until the correct one was discovered. The subject learns nothing about the solution to the problem before he attends to the correct cue, except that his presolution hypotheses were wrong. For example, when a rat is operating on the hypothesis that left is the correct cue, he is learning nothing about the correctness of the black and white cues.

Continuity theory (Spence, 1936) was conceived as an alternative to the noncontinuity theory. This theory assumed that learning was continuous in the sense that every response to the correct stimulus (for example, black) increased to some degree the tendency to approach the black alley, whereas every response to the incorrect stimulus (for example, white) reduced to some degree, the tendency to approach the white alley. Spence also assumed that the learning was nonselective in the sense that, on each

rewarded trial the strength of the tendencies to approach each component of the stimulus compound was increased, and decreased when nonrewarded. Thus, for example, when black is chosen on the right side, the tendency to approach black and right is strengthened. Since only the tendency to approach black is consistently strengthened, it gradually attains ascendency and finally becomes sufficiently strong so that the subject chooses the black alley most of the time.

The experimental design used to settle the continuity–noncontinuity dispute was as elegant as it was simple (Spence, 1940). During the presolution period, when the subject is responding at a chance level, the reinforcement contingencies are reversed. That is, for our sample problem, *white* is now reinforced, whereas *black* is consistently not reinforced. According to the noncontinuity theory, this reversal shift (reversing the reinforcement contingencies) should not retard learning. If responding on a chance level to the relevant cues indicates that the subject is attending to an irrelevant cue, usually position, then shifting the reinforcement contingencies of the relevant cues should have no effect on subsequent learning. Continuity theory generates an opposite conclusion: a reversal shift should retard problem solution. Even though the subject is responding at a chance level during the presolution period, the correct habit (response to black) is gradually strengthened whereas the incorrect habit (response to white) is gradually weakened. Reversing the reinforcement contingencies would retard learning because the effects of prior learning would now have to be overcome. Continuity theory, but not noncontinuity theory, would predict that a control group with unchanged reinforcement contingencies should solve the discrimination task more rapidly than a reversal-shift group.

In reviewing the continuity–noncontinuity controversy Deese and Hulse (1967) conclude:

> There are a number of experiments which have used the technique of reversing cues during the presolution period, and the overwhelming majority show that reversing cues *does* retard learning. . . . It is clear, then, that the results of the discrimination experiments in which cues are reversed early in learning favor the continuity view [p. 190].

Riley (1968) drew a similar conclusion: ". . . the evidence predominantly supported the continuity position [p. 121]."

In spite of these successes, Spence, as already noted (Section I.B.2), emphasized the boundary conditions of his formulation by noting that it was intended to apply to nonarticulate organisms. His student (Kuenne, 1946) demonstrated that continuity theory could predict the discrimination learning of "preverbal" children but not account for the behavior of "verbal" children. This study anticipated our own investigations of phylogenetic and ontogenetic differences in discrimination learning.

B. Discrimination-Shift Behavior in Humans and Animals

Much of the research to be cited is based on some variant of the mandatory discrimination-shift paradigm illustrated in Figure 1. The subject first learns a preshift discrimination between two discriminanda that differ simultaneously on two dimensions, for example, brightness and shape, but only one dimension (for example, *brightness*) is relevant to the reinforcement contingencies (for example, *black* is correct and *white* is not). The other dimension (*shape*) is irrelevant because neither *circle* nor *triangle* is correlated with reinforcement. After learning the initial discrimination the subject encounters one of two possible postshift discriminations: a reversal (RS) or extradimensional shift (EDS). In a RS the reinforcement contingencies in the postshift problem are reversed from the preshift problem (for example, *white* is now correct and *black* is incorrect). In an EDS the previously irrelevant dimension (*shape*) becomes relevant with one of its cues correct (for example, *circle*) and the other (*triangle*) incorrect.

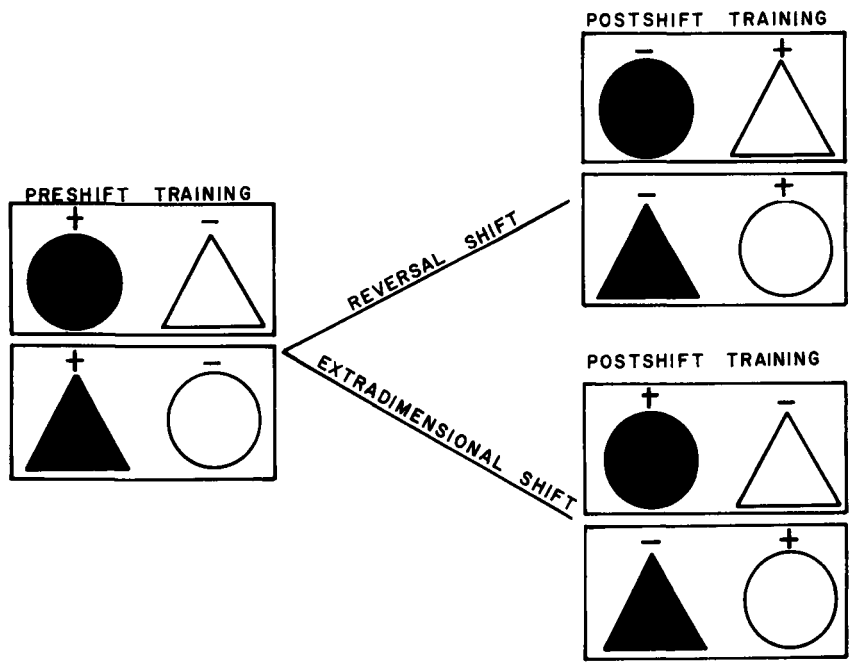

FIGURE 1 A mandatory discrimination-shift paradigm involving a comparison between a reversal (RS) and extradimensional shift (EDS). For each pair of discriminanda, plus indicates reinforcement; minus indicates nonreinforcement. A counterbalanced design is conventionally used so that each cue (*black*, *white*, *circle*, and *triangle*) is correct for one-fourth of the subjects during both the preshift and postshift problems.

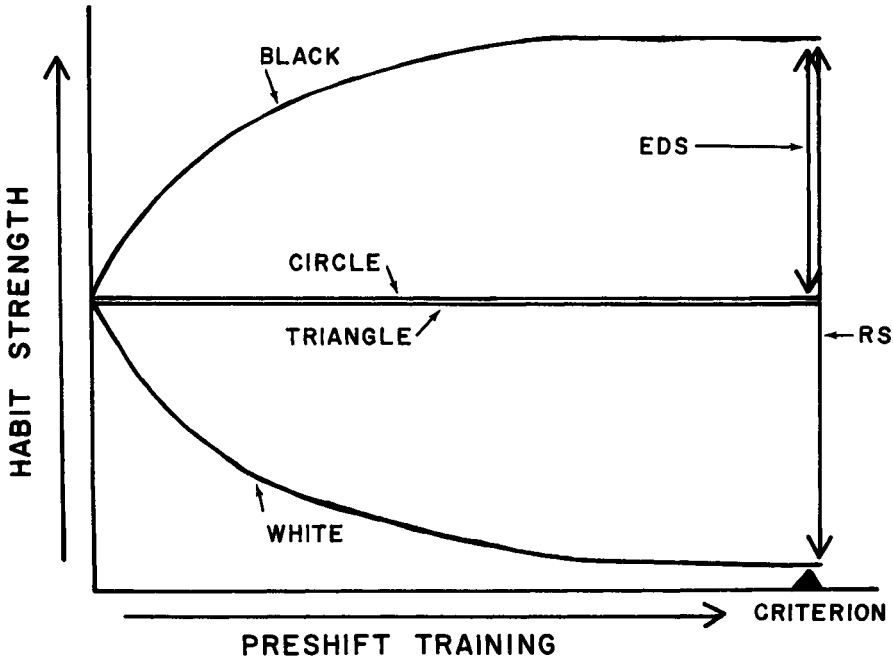

FIGURE 2 Schematic representation of difference between competing habits at end of preshift training. During preshift training, the positive stimulus, *black,* is reinforced consistently thus increasing the tendency ("habit strength") for the subject to approach it. White, consistently nonreinforced, loses habit strength. The irrelevant cues, *circle* and *triangle,* intermittently reinforced are represented as maintaining their initial moderate level. The significant feature of this theoretical representation is that at the end of initial training the difference between the habit to approach *black* and either of the two irrelevant cues, *circle* and *triangle,* is less than the difference to approach *black* and the incorrect cue, *white.* Because of this, more training would be required to execute a RS, from *black* to *white,* than an EDS, from *black* to one of the previously irrelevant cues, for example, *circle.*

Continuity theory (Spence, 1936) would deduce that, in a counterbalanced design, a RS is slower than an EDS (H. H. Kendler, Hirschberg, & Wolford, 1971). The fundamentals of this deduction are illustrated in a highly idealized form in Figure 2. College students were found, however, to execute a mandatory RS faster than an EDS (Buss, 1956). H. H. Kendler and D'Amato (1955) in interpreting this finding concluded that the Spence model is inappropriate for the discrimination-shift behavior of college students, a judgment consistent with Spence's own emphasis on boundary conditions of the single-unit model: "I have long since lost track of how many disproofs there have been of my theory of discrimination learning in nonarticulate organisms that have employed college sophomores as subjects" (Spence, 1956, p. 202).

A mediational S–R model was offered (H. H. Kendler & D'Amato, 1955; H. H. Kendler & T. S. Kendler, 1962; H. H. Kendler & Vineberg, 1954) to account for the discrimination-shift behavior of the college students. In the single-unit model of Spence's original discrimination-learning theory the input component, the stimulus, is directly linked, in a figurative sense, to the output component, the response. This model predicts the choice responses of subjects from the physical characteristics of the stimulus components and the number of pairings they had with reinforced and nonreinforced responses. In contrast, the mediational model involves an additional theoretical link, a mediational mechanism that intervenes between the observable stimulus and response, operating to transform incoming stimulation into some internal representation that guides subsequent behavior.

The implications of these two models for discrimination-shift behavior were initially represented by H. H. Kendler and T. S. Kendler (1962) in the schematic form of Figure 3. Unlike the single-unit model, the mediational model predicted a more rapid RS than EDS because a RS involved a common mediational link during both pre- and postshift training, while an EDS required discarding of the preshift mediational sequence and the substitution of a new one. Without doubt, this theoretical analysis contained elements of ambiguity. But ambiguity in the early stages of theoretical development, when empirical evidence is sparse and many theoretical alternatives are possible, may be a more productive strategy than premature precision that is overcommitted to one of several reasonable conceptual

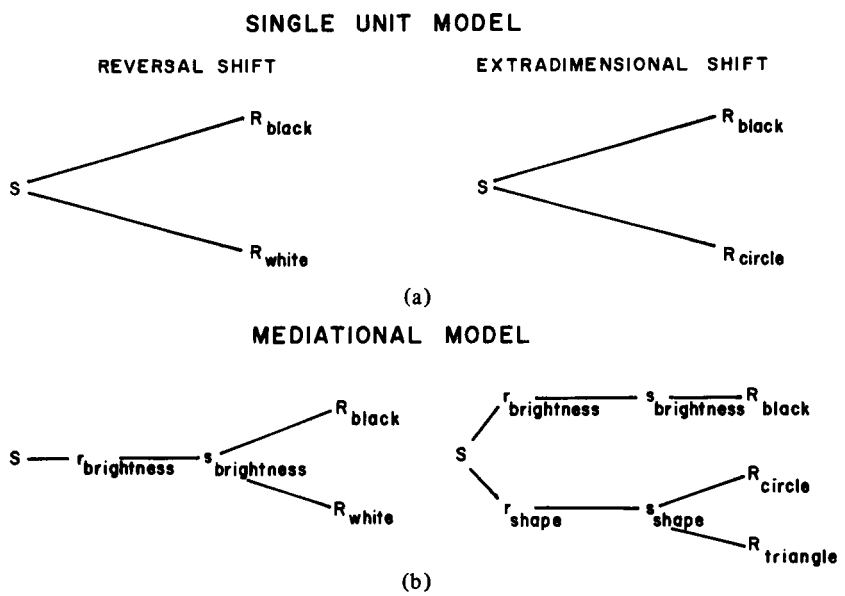

FIGURE 3 A single-unit (a) and mediational (b) S–R analysis of reversal and extradimensional shifts.

options. During this early period when the major emphasis was on the heuristic value of the coordinated single-unit and mediational approach to discrimination-shift behavior, the following two orienting attitudes operated:

1. The crux of the distinction between the single-unit and mediational models is that in the former, behavior is directly connected to *specific* attributes of physically defined discriminanda, whereas in the latter the behavior is correlated to the subject's *conceptual transformation* of the discriminanda. In other words, the discriminanda in the single-unit mechanism are "processed" as discrete stimuli while the mediational mechanism "processes" the discriminanda as instances of a conceptual category.
2. The exact nature of the mediational process was left open ended because of the reasonable expectation that several distinct mechanisms would be involved (H. H. Kendler & T. S. Kendler, 1962; T. S. Kendler, H. H. Kendler, & Learnard, 1962). Mediation was not conceived as a single internal mechanism but instead as a set of different mechanisms, all of which transformed in some manner incoming stimulation. The conceptual properties and operating characteristics of these different mediational mechanisms could presumably be inferred from experimental results.

The fact that human adults behaved according to the mediational formulation in the RS–EDS comparison immediately raised the question of whether infrahuman subjects would respond in a manner consistent with Spence's single-unit model (1936, 1937a). Kelleher (1956) answered this question affirmatively when he found that rats, when trained to a criterion, executed an EDS more rapidly than a RS. Other investigators, sometimes with variations of the basic mandatory RS–EDS design, confirmed Kelleher's findings by reporting evidence that in addition to rats; fish, chickens, pigeons, and young and old monkeys found an EDS easier to execute than an RS (Brookshire, Warren, & Ball, 1961; Schade & Bitterman, 1966; Tighe, 1964).

Research issues. The discontinuity between the behavior of adult humans and infrahuman subjects suggested several lines of research. One obvious one was the developmental analysis of human discrimination-shift behavior, which was expected to throw light upon the origins and operations of mediational mechanisms. In planning, executing, and interpreting such research we attempted to minimize unnecessary semantic confusions and false theoretical dichotomies generated by the perennial controversies about the relative merits of S–R and cognitive models of learning and the relative importance of learning and perceptual processes (H. H. Kendler, 1971; H. H. Kendler & T. S. Kendler, 1962). We even hoped that our