

The meaning of representation in animal memory

H. L. Roitblat

Department of Psychology, Columbia University, New York, N.Y. 10027

Abstract: A representation is a remnant of previous experience that allows that experience to affect later behavior. This paper develops a metatheoretical view of representation and applies it to issues concerning representation in animals. To describe a representational system one must specify the following: the *domain* or range of situations in the represented world to which the system applies; the *content* or set of features encoded and preserved by the system; the *code* or transformational rules relating features of the representation to the corresponding features of the represented world; the *medium*, or the representation's physical instantiation; and the *dynamics*, or how the system changes with time. In part because of the behaviorist assumption that the hypothetical, covert changes occurring in an organism during learning correspond to the overt physical changes that are observed, issues of representation in animal behavior have been largely ignored as irrelevant or misleading. However, it can be inferred that representations, acting as models of environmental regularities, operate at many levels of behavioral functioning, both cognitive and noncognitive. Objections to the use of this concept in explanations of animal behavior, based on the claim that it is indeterminate and on behaviorist considerations of parsimony, can be answered. Animal representations may be specialized in terms of tasks and species. Data from tasks involving spatial memory, delayed matching-to-sample, and sequence learning suggest some foundations for a general theory of animal representations.

Keywords: animal cognition; coding; cognitive maps; imagery; memory; monkeys; pigeons; rats; representation

How the results of experience are maintained or “represented,” so as to affect later behavior has been one of psychology’s longstanding and central concerns. The meaning of the concept of representation has not been well understood, however. This paper extends the metatheoretical view of representation that Palmer (1978) has suggested (taking a broader and, one hopes, more solid view of representation than is typical) and applies it to issues concerning representation in animals.

A metatheory of representation

In order to describe how animals represent the effects of experience, I will first approach the problem of representation in more general terms. A representation (or, more properly, a representational system) requires at least two “worlds,” or sets, in which the events, objects, or features (henceforth, collectively, *features*) of one world “represent” some (but not necessarily all) features of the other. A representation is a kind of model of the represented world; particular features of the representation “stand for” corresponding features of the represented world. The concept of representation calls for a mapping relation between the features of the represented world and those of the representation (though this correspondence need not be isomorphic, or one-to-one). The representation preserves information about the represented world to the extent that it preserves corresponding relations among the features of the two worlds.

Representations can differ along several dimensions, representing the same information in different ways. A complete description, or theory, of a representational

system requires the specification of its domain, content, code, medium and dynamics.

The *domain* of a representational system is the circumscribed class of situations or tasks in which it is used and to which it applies. Different features (different distinctions) are important in different tasks. A representational system encoding one particular set of features for one task may be wholly inadequate for another task. A representation’s domain delimits its applicability.

The *content* of a representation consists of those particular features of the represented world that are preserved in the representation, determining the kinds of information that can be derived from it. Two representations are informationally equivalent if the same features of the represented world are preserved by both.

The *code* of a representation consists of the transformation rules underlying the mapping between the features of the representation and the corresponding features of the represented world. Palmer (1978), for example, describes seven different codes that could be used to represent a few features of a set of rectangles. Each of the codes contains information reflecting (i.e. standing for) some features of the represented world. The retrievability of that information depends on the transformational code (and its invertibility).

The *medium* of a representation is its physical instantiation. The medium itself does not provide information about the represented world; it merely supports (the way paper supports ink, or hardware supports software) the information-carrying aspects of the representation. The physical form of a weather report, for example, can change from one medium to another (from a written message on teletype paper, to a spoken report, to radio

waves, to sound, etc.) while preserving the same information. Ultimately, some of the information in the weather report, such as the expected temperature and the likelihood of rain, may be "represented" in the listener's behavior, coded in terms of the type of clothing worn or the presence or absence of an umbrella. As the medium changes, the code may also change (sound patterns to modulations of radio waves); or the code may not change (vibrations of the sound waves from speech to vibrations of the sound waves in the microphone), even though the medium does change. (Hence, the medium is not the message.)

The dynamics of a representation pertain to its changes over time. Of particular interest in most studies of representational systems is the manner in which a representation comes into being, and how it changes (and, in some systems, ceases to be available). Also of interest are the changes occurring over time in the information retrieved from a representation (in cognition this may be due to such factors as incomplete concept formation or partial forgetting, for example).

A representational system includes *processing* components that determine the relations among features of the representation, and interpret them as symbolic counterparts of the features of the represented world. The relations among features of a representation may have no resemblance to the corresponding features of the represented world except as dictated by the (potentially arbitrary) correspondence rules of the code. Therefore, representational systems must include processes embodying those rules. These processes are a necessary part of the representational system, and, in some cases, may themselves be encoded in it (the way a computer program may be encoded in the same memory system as the data on which it operates).

A representational system's computational or processing characteristics can provide valuable information about its other properties. For example, some codes may allow one type of task to be performed rapidly, but require more time for other tasks. This predicted difference can then be used as a basis for comparing the two alternatives. Two representations are *computationally equivalent*, on the other hand, if the same information can be extracted from each with about the same amount of computation (Simon 1978).

Representations and memory

As cognitive scientists, we investigate the role played by representations in the control of behavior; at the most general level, we are interested in those representations that allow past experience to affect later behavior. After an experience has ceased, some change must endure to mediate subsequent effects of that experience: That change, whatever its form or substance, is what is meant in the rest of this paper by a representation.

Representation is not synonymous with memory. When we speak of memory we typically mean representations internally encoded in the organism's nervous system; but representations in external media (such as books) can also control behavior. In this paper I will call representations in a neural medium "cognitive represen-

tations." This is not to say that external, nonneural representations that affect behavior do not require cognitive processing, but only that such external representations are not themselves cognitive.

One example of a noncognitive representation can be seen in the foraging behavior of mantids. Theories of foraging strategies assume that individual predators have a representation of the rate at which they have been obtaining food (Charnov 1976a). In some animals this information is undoubtedly represented neurally (cognitively). In other cases, however, a more direct representation has been found. Mantids were found to represent their feeding rate by the volume of their gut contents. Based on that information, and only indirectly on the rate at which they had been feeding, they were found to make appropriate and optimal foraging decisions in attacking flies placed at various distances (Charnov 1976b).

In cognitive science it is so far more convenient to avoid being too concrete about the physical instantiation of cognitive representations. We concentrate instead on the nervous system as a "virtual machine" (Pylyshyn 1980), analyzed not at the level of individual neurons and their structure, but at that of operations and contents. Similarly, it is often necessary to describe cognitive codes by analogy, because we do not, as a rule, have access to the actual codes. For example, we are usually limited to describing codes by analogy to properties of observable stimuli and responses. Unless this is done carefully, however, it leads to confusion between the information being stored (content) and the means by which it is stored (code). An "image" is an example of the kind of analogical description typically used to specify a representational code. No one is likely to argue that images are actually pictures in the head, yet it does seem reasonable to describe images in terms of their picture-like properties (Kosslyn 1981).

Working and reference memory. In many tasks used to study animal memory, the subjects are required to base their performance on stimuli that are no longer present. For example, in delayed matching-to-sample, animals (typically pigeons) are presented with a sample stimulus for a time period, and then, following a retention interval, with a number of alternative stimuli. Subjects are rewarded for choosing the matching comparison stimulus. To choose the correct alternative, the animal must, during the delay, remember information about the identity of the sample stimulus. This information is held in "working memory" (Honig 1978). In general, working memory retains information about those particular stimuli occurring on a trial that are important in determining appropriate actions. Because different sample stimuli are presented on each trial, the information contained in working memory is typically of no use to (and can interfere with) later performance; hence the optimal strategy is to "flush" or forget the contents of working memory at the end of each trial.

In addition to the information held in working memory, certain transformational rules are needed to specify the relationships between samples and correct choices. This mapping relation is constant from day to day and could be stored in "reference memory" (Honig 1978).

Reference memory contains the subject's knowledge base concerning the relationships among stimuli, the outcome of particular stimulus-behavior combinations, mapping relations between stimuli and behavior, etc. Both kinds of information are necessary for adequate performance (Honig 1978).

Objections to the concept of representation

To assume the existence of a representation is rather innocuous and should rarely be an issue for theoretical dispute. If an organism's behavior can be shown empirically to be affected by past experience, then some representation of that experience *must* exist. Simply positing the existence of a representation, however, explains nothing; the nature of that particular representation, rather than the fact of its existence, is what provides a kernel of an explanation, and is the proper target for dispute. This issue has not been well-understood, and has as a result been a major source of resistance to the use of the concept of representation and animal memory.

Behaviorists' approach to representation

One characteristic behavioristic (and neobehavioristic) objection to the use of the concept of representation in explaining animal behavior (Segal & Lachman 1972; Weimer & Palermo 1973) is that the concept is at best irrelevant, at worst distracting or misleading. Behaviorism grew, in part, as a reaction against the mentalism of introspectionist structuralism (Marx & Hillix 1979). For the behaviorist, any concepts that even hinted at subjectivity, or unobservability, such as mind and its contents, were considered inappropriate subjects for scientific inquiry. Behavior was to be understood as a function of the relations between stimulus events and response outputs. Specifying the lawful relations among these observables (stimulus-response associations), without addressing the mechanism mediating those relationships, was thought to be sufficient by way of explanation; only objectively observable or logically (i.e. deductively) derivable entities were permissible (Weimer & Palermo 1973).

The correspondence assumption. The behaviorist solution to the problem of representation was to assume that psychological processes underlying an observed behavioral change were isomorphic to, or in correspondence with, situational and behavioral variables (Bolles 1975). For example, Hull asserted that the acquisition of knowledge was equivalent to the acquisition of habits, which was synonymous with the acquisition of new reflex arcs through experience (Hull 1943). Changes in the conditional probability of a behavior, given some stimulus, were assumed to be a result of, and to operationally define, a strengthened connection between that stimulus and that response (S-R or stimulus-response associations). The correspondence assumption made it possible to behave as if the observed variables *were* the theoretical variables, and so to ignore issues of representation.

In later work, if the concept of representation was considered at all, it was as an alternative to S-R (stimulus-response) based explanations. The question became

whether a given behavior was a result of a representation, or if it constituted a set of reflex-like S-R connections. Such a question is based on a misconception. Representation is not offered as an alternative to S-R models; explanations in terms of S-R associations are simply one possible variety of representational system that can be understood according to the representational metatheory presented earlier.

S-R representations. The domain of the S-R theories was thought by behaviorists to be practically unlimited. Under the influence of the theory of evolution and its implications for interspecific continuity of behavioral mechanisms, most behaviorists assumed that all complex psychological processes were reducible to the same mechanisms present in nonhuman animals. Laws of learning could be studied in any particular animal and the results generalized to most other organisms.

The content of S-R representational systems consisted exclusively of associations between stimuli and responses. In early theories (e.g., Thorndike, 1911), no other information was thought to be present. Experiments demonstrated, however, that animals were encoding more information than straightforward S-R connections (Maier 1929; Tolman 1932). Behavioral contrast experiments (e.g., Flaherty & Largen 1975) indicate that animals have access to information about reinforcer magnitudes (see also, Hulse 1978). The presence of this additional information was accommodated through the inclusion of covert stimuli and responses in the S-R formulation. Covert stimuli, perhaps arising from kinesthetic feedback from prior responses or from other covert responses such as fractional anticipatory goal responses, were thought to become associated with, and therefore elicit, those behaviors apparently indicating, say anticipation (Hull 1943).

The medium and code of the S-R representational systems were not usually considered, except to the extent that they were thought to involve minute copies of overt behavior in the form of small-muscle movements, such as in subvocal speech (Watson 1920).

The dynamics of S-R systems was one area that did receive much attention. The search for the "true" form of learning curves, for example, motivated a considerable amount of research. Continuity theorists held that learning occurred as a gradual build-up of association (Hull 1943; Spence 1937). Noncontinuity theorists held that learning was a result of sudden, all-or-none shifts (Krechevsky 1932; Lashley 1929). Neither position received universal acceptance, but both persist to the present day (Riley 1968).

A great deal of evidence, accumulated from diverse sources, indicates that the correspondence assumption, or any similar assumption that learning consists exclusively of the formation of connections between stimuli and responses, is inadequate to explain animal and human behavior. There is not enough space here to recapitulate the arguments leading to this assertion, but they are well-documented elsewhere (e.g. Bolles 1975). Their most important features are (a) that more aspects of a situation are represented than simple reflex-like S-R connections, and (b) that the relationships between situations and behaviors contain more structure than could

easily be accommodated by S-R associations (Gallistel 1980; Oatley 1978). These points are discussed below.

The problem of indeterminacy

A second objection to the use of the concept of representation arises from the unobservability of representational systems. Although the testability of a theory does not depend on a one-to-one correspondence, or isomorphism, between conceptual and observable events (Estes 1976), observable phenomena do remain the only available basis for making such tests. Inferences about the covert functions of an organism must still be based on observable phenomena. The problem involves the ability to test alternative models of those internal processes solely on the basis of observable data.

This problem of indeterminacy has been most clearly presented in the controversy over whether it is necessary to consider picture-like images as a code for human memory, or whether complex propositional systems are sufficient to account for all human memory representations (Anderson 1978; Hayes-Roth, 1979; Pylyshyn 1979). Anderson has argued that:

It is not possible for behavioral data to uniquely decide questions about a representation. . . . One cannot just test questions about a representation in the abstract. One must perform tests of the representation in combination with certain assumptions about the processes that use the representation. . . . One must test a representation-process pair. Given any representation-process pair it is possible to construct other pairs with different representations whose behavior is equivalent to it (Anderson 1978, pp. 262-3).

The basis of Anderson's argument is the assumption that for any particular representational system (consisting of representations and processes) one can prove the existence of *at least one* other representational system involving different representations and processes that makes *identical predictions* in all respects and so cannot be discriminated from the first *on any empirical basis*. This claim is potentially devastating to the study of representation in humans as well as in animals. If no empirical basis can be sufficient to help one choose among alternative theories of representational systems, then there is little point in empirical investigations.

Careful examination of Anderson's arguments in the light of the metatheory described here suggests that their potential for devastation is more apparent than real. In order for two representational systems to mimic each other (i.e. provide identical predictions), they must meet stringent requirements of equivalence. First, they must be informationally equivalent, representing the same information, and preserving equivalent distinctions with equivalent resolution. If one model preserved a distinction not present in the other model, then that distinction could subserve a behavioral difference that would provide a basis for choosing between them.

Second, they must be computationally equivalent (Simon 1978; Pylyshyn 1980, has called this "complexity-equivalent"). They must both produce the same output with the same amount of computation in every situation. The two models need not involve the same number of steps to be computationally equivalent, but for any particular input, the number of steps, or complexity, for one

model must be linearly related to the complexity of the second model.

Third, Anderson's proof (Anderson 1978) relies on a peculiarly constrained alternative in which there is a one-to-one isomorphism between the codes of one representation and the codes of the mimicking representation. At least one such isomorphic representation must exist if the inverse of the coding process is computable. Two models are codally equivalent if their codes are isomorphic, part for part, with each other.

The two mimicking representational systems on which Anderson's proof is based are *informationally equivalent*, preserving the same distinctions about the represented world, *computationally equivalent*, taking equivalent amounts of time to perform each task, and *codally equivalent*, having codes that are isomorphic with one another. The two systems not only represent the same information about the represented world, but also, because of their equivalences, represent each other. That is, Anderson's mimicking model exists only when it is a representation of the model it mimics, differing only, perhaps, in the names assigned to the representations, codes, processes, or states. In all other ways they are identical, and the problem turns out to be a pseudoproblem.

If, on the other hand, informational and computational equivalence can be maintained without codal equivalence, then the problem of indeterminacy remains serious. However, this has not yet been shown to be possible. Hayes-Roth (1979) and Pylyshyn (1979) outline a series of other objections to Anderson's (1978) claims, dealing more particularly with the distinction between image- and proposition-based representations.

Early investigations of representation in animals

Hunter. Acceptance of the correspondence assumption was not universal among students of animal learning and behavior. For example, Hunter (1913) trained animals in a delayed response task to observe a light over one of three doors behind which food was hidden. The animals were then confined in a start box until some time after the light went off, then released and allowed to open one door. "If a selective response has been initiated and controlled by a certain stimulus, and if the response can still be made successfully in the absence of that stimulus, then the subject must be using something that functions for the stimulus in initiating and finding the correct response." (Hunter 1913, p. 2)

Hunter was reluctant to assume that so-called higher mental faculties were at work in the animals that he tested, so he considered three alternatives that might account for their performance. The most obvious alternative Hunter considered was that of "overt orienting attitudes" in which the animal would identify the correct compartment and represent that information by positioning itself in a particular manner, usually oriented toward the correct door (cf. Blough 1959; Cohen, Looney, Brady & Acuella 1976). A second form of representation considered by Hunter, one he thought was used by raccoons and children, was "intra-organic cues" or "sensory thought," which does not include the use of imagery. Hunter considered the highest form of representation to be thought that included imagery. We may now question the

validity of his classes of representation, but the point is clear. The animals were using something to stand for the information about the correct door.

Hunter continued to investigate the concept of representation in subsequent work with a "temporal maze" (Hunter 1920) in the form of a squared-off figure 8. Animals were required to perform a sequence of responses such as two loops from the central arm of the maze around the left block followed by two loops to the right. The correct choice at the end of any given loop depended on the choices made on the two prior loops. Following the first left loop another left loop was correct, but following the second left loop a right loop was correct. In Hunter's view, correct performance on this task was analogous to human counting. Whatever cues were used by the animal were internally generated as the maze did not change from one choice to the next. Hunter called these internally-generated cues representations.

Lawrence. Lawrence (1963) noted a parallel between problems of representation and problems usually associated with psychophysics. In both fields an external event called a stimulus is manipulated by an experimenter and corresponding changes in behavior are noted. One of the central assumptions of psychophysics is that the proximal stimulus, as the organism perceives it, is not identical with the distal stimulus, as a physicist might measure it. On the basis of this distinction Lawrence formulated a concept of representation, the *stimulus-as-coded*, based on a conceptual mechanism called the *coding response* that mediates between the proximal stimulus and the response, and codes the stimulus. Lawrence described it this way:

By coding the following is meant: If there is a set of objects or events and to each of them a different label is assigned, then the labels code these objects or events. More generally whenever a one-to-one correspondence between two sets of events is established, either set of events can be taken as a code representing the other set. Nothing more than this is implied by the term coding as used in the present context (Lawrence 1963, p. 188).

Tolman. Tolman (1948) organized his ideas on representation around the concept of the cognitive map. The basic idea behind a cognitive map is that the animal organizes its behavior, not in terms of simple associations between stimuli and responses, but rather as a complex organization of the potential goals and the means of reaching those goals (means-ends expectancies). The mechanism whereby cognitive maps were thought to arise was the formation of stimulus-stimulus (S-S) associations.

Characteristics of representation

Varieties of representational systems. Representations can be seen to operate at many levels of behavioral functioning. The simplest homeostatic mechanism, for example, contains a kind of primitive representation in the form of a reference level, or set point, representing the preferred state of that mechanism (Oatley 1978), and a deviation from the set point, representing recent experience. Behavioral strategies based on such simple systems can react only to recent situations. Anticipatory systems

that can predict or avoid unwanted states require more elaborate representations. In behavior theories, such as Hull's, these anticipatory processes took the form of mechanisms based on interoceptive covert responses and stimuli (r_g-s_g) similar to Tolman's means-end expectancies. Coupling the theories in reflex-like stimulus-response terms allowed behaviorists to ignore issues of representation, but it did not resolve those issues.

Models

As Tolman (1932, 1948) made clear, means-end expectancies, whether described as covert associations or as significant Gestalten, can be organized into cognitive maps or models of a situation. The use of models as controllers of behavior, instead of, or in addition to, immediately experienced stimulation is quite ubiquitous, playing a role in a large number of behavioral systems. Circadian oscillators provide one example of this sort (Oatley 1978). Because the rhythms can be entrained to light-dark cycles as experienced, they clearly mediate the effects of that previous experience on later behavior. Their importance in the present context, however, stems from their role as models of the earth's rotation, or of the light-dark cycle. Circadian oscillators function as models of the environment because they share common properties (i.e. oscillation) with those light levels. Neurobehavioral processes such as sleep are controlled by these models, not by direct stimulation (Oatley 1978). As models, they allow the animal to take advantage of the regularities in the environment without a constant goad, permitting the organism to fill in parts of an immediate or recently past situation that were not directly experienced on that occasion. They also, thereby, insulate the organism from unimportant random fluctuations, and provide a mechanism by which the organism can anticipate relevant changes.

Maier's reasoning experiments. Similar use of a representation to fill in blanks missing from direct experience can be seen in the detour behavior of rats studied by Maier (1929). Maier ran three rats through complex, three-dimensional mazes made up of tables, ringstands with wire mesh ladders, and narrow plank runways. During the first part of training this set-up was absent from the room (though one or more tables may have been present) and the rats were allowed to explore freely. During the second part of training, another table (Table C) was added (see Figure 1). This new table was connected to another table (Table A) by a narrow plank of a sort with which the rats were familiar and on which they had been trained to run; it was also connected, by similar planks, to each of three ringstand ladders which the rats had been trained to climb. A fourth ringstand was accessible from Table A. During this phase of training, food was placed on one corner of Table A, accessible via the plank from Table C, but separated from the rest of Table A by a cage. Through a series of three training trials in which they were started progressively farther from the food, the rats were trained to run from a point on the floor to a particular ladder, up the ladder to Table C, and from there to the food.

Following this training the rats were placed on Table A, on the side of the barrier opposite the food, but with

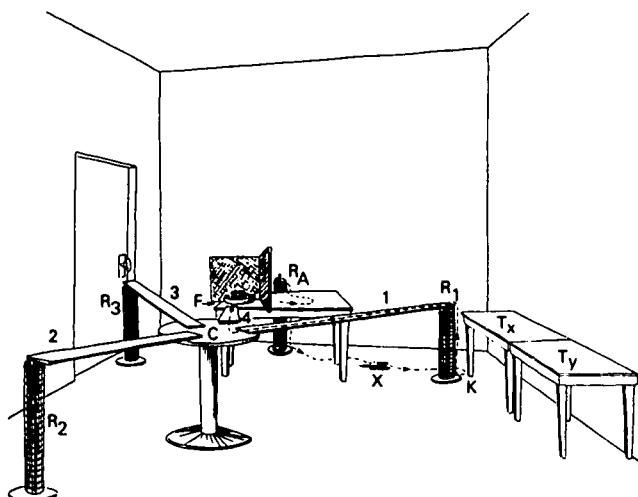


Figure 1. The apparatus used by Maier (1929). Essentially the same set-up was used in both the familiar and the novel rooms. Reprinted, with permission, from Gallistel (1980).

access to the ringstand ladder. After running about on Table A, the rats ran down the ringstand ladder and directly to the ringstand they had been trained to climb during the second phase of the experiment; they climbed it to Table C and went on to the food. This experiment was repeated with three further rearrangements of the same apparatus, then with similar arrangements in a novel room that the rats had not explored.

Two differences between the performance in the familiar room and novel room were apparent. First, the rats spent much more time on the floor in the novel room than in the familiar one. Second, in the familiar room the rats always climbed the ladder they had been trained to climb during the second phase of the experiment and never any of the other ladders. (At least one rat was trained to run from the base of one ladder, across the floor, and up a second ladder.) In contrast, the same rats in the novel room chose ladders at random on their first test run from Table A, showing no tendency to climb the ladder on which they had been trained. (On two subsequent runs in the same set-up, the likelihood of choosing the original ladder did increase.)

These findings could be attributed to any of three potential explanations. First, performance in the novel room could have differed from performance in the familiar room because of generalization decrement. The second room may have been sufficiently different from the one in which the rat had experience to disrupt performance. Second, the difference in the two tasks may represent response competition from exploration of the novel room. Third, the difference could be attributed to the lack of an adequate representation (e.g. a cognitive map) for the novel environment, which is a necessary foundation for the additional information derived from the second training phase.

Generalization decrement is not an explanation, but simply a restatement of the experiment's result: It is merely an assertion that an experiment presents an example of a familiar phenomenon. Also, the rats did have experience in the novel room, consisting of four training trials (versus three in the familiar room) in which they were trained to run to the food from various parts of the

set-up and from various locations in the room. The role that transfer might play in the novel room is therefore uncertain.

The second explanation, response competition, could account for the increased time the rats spent on the floor in the novel room, but it fails to explain why the rats failed to climb the familiar ladder when they did eventually climb a ladder. It is unclear how competition not only delayed traversal of the familiar path, but also induced the rats to choose their ultimate path by apparent trial and error.

The third explanation, inadequate environmental representation, may be the most reasonable. Some kind of experience (e.g. exposure to the room, exposure to that part of the floor, etc.) with the points on the floor between the ringstand next to Table A and the ringstand the rats were trained to climb was crucial to efficient traversal from one to the other. Furthermore, as Gallistel (1980) and Maier (1929) have argued, it is difficult to attribute the efficient performance of the animals in the familiar room to learning or habit formation because the particular combination of responses that got the animal from the base of the ringstand at Table A to the trained path had probably occurred no more often than any other combination of responses. The rats were not using the sight of the familiar ringstand or any other associated factor to control approach to it; in the unfamiliar room, some of the rats wandered near the familiar ringstand ladder before climbing a different ladder. Finally, the rats were exposed, on each day of the experiment, to a different arrangement of tables, ringstands and planks; yet performance in the familiar room was always efficient and the ringstand on which they were trained that day was always the one chosen, further strengthening the claim that general knowledge about the room was the main factor producing the differences between the novel and familiar room.

Animal navigation. Other experiments on animal navigation also suggest that animals use experience-derived models of their environment to control behavior. For example, displacement studies with migrating starlings (e.g., Perdeck 1958) show that experienced birds use some kind of map of their migratory route (perhaps referred to solar or stellar coordinates) to control their flight path. Perdeck captured starlings in Holland during their autumnal migration from Eastern Europe and displaced them by air to Switzerland, where they were released. Adult starlings compensated for this displacement and flew in a northwesterly direction, ending in their usual wintering ground. Juveniles (birds who had never migrated before), in contrast, did not compensate for the displacement and continued in the same direction they had been flying before capture, ending their migration in an area where others from their breeding population were not found (though starlings from other summering locations were found in that area), and continuing to winter there in subsequent years.

Similar tolerance to displacement has been found in insects such as the digger wasp. The female wasp lays eggs in small burrows dug into the ground. In the process of provisioning the nest she makes a stereotyped reconnaissance flight over the area, and departs. The displacement of conspicuous landmarks (twigs, pine cones, etc.)

from around the nest by a few meters (Tinbergen & Kruyt 1938) causes the wasp to search for the nest in a position appropriate to the displacement, away from its actual location, indicating that the approach to the nest from its immediate vicinity is controlled by these particular landmarks (whose locations were presumably learned during the reconnaissance flight). On the other hand, displacing the wasp, by moving her in a dark box over distances as great as 50 meters from the nest, fails to interfere with her return to the nest (Thorpe 1950). The significance of these studies is that no set of habits, or beacons, could allow the animals to solve these problems as they do. They involve novel responses and appear to require the presence of a representation of the sort provided by a cognitive map operating as a model of the situation.

A broad range of phenomena, ranging from circadian rhythms to navigation, have been considered here under the rubric of modelling representations. They all share features that justify their joint consideration. Each involves a representation of environmental regularities, derived from experience and controlling behavior. There is symbolic representation of goals, desired states, or meaningfully structured activities in relation to some aspect of the world. Their differences lie in their level of operation. All animals that have been tested (and probably all plants), possess circadian oscillators. The other representations are more uniquely specialized to the particular species possessing them. In including such a broad range of modelling representations, my intention is to show that there is nothing mystical about the existence of representations, and to illustrate the kinds of selective advantages leading to the evolution of representational mechanisms. Our goal as psychologists is to discover the features of the particular representational system that affect or control the behavior we have chosen to study.

The scope of representational systems

Aside from the problems associated with inferring covert processes, the problem of studying representational systems in animals is further complicated because a given representational system may not be universal to all species in a given situation, or to all situations for a given species. The particular representation formed is likely to depend on the task at hand and on the species. Von Uexkull (1957) argued that each animal lives in a species-characteristic sensory world. As a result, we might expect each animal to encode different characteristics of that world in different types of representations, depending, at least in part, on the different perceptual mechanisms and abilities of each animal. These representations would be further specialized to the task. Some types of representation are better suited for solving some kinds of problems than are others. In human problem-solving, for example, the way a problem is represented can make a crucial difference in how successfully it is solved (e.g. see Oatley 1978, pp. 141–147).

Formal properties and task dependent representations. Representations can also be task-dependent, in the sense that not all tasks may make use of the information contained in a representation. One example of this kind of specialization is seen in the case of training win-shift vs. win-stay discriminations (Olton, Handelmann & Walker

1981). In win-stay discriminations, animals are reinforced for returning to a location where they have previously been fed. In win-shift discriminations, subjects are reinforced for going to a location different from that in which they had just been fed. Experiments of this type have been performed with monkeys, rats, and fish. In my laboratory we trained Siamese fighting fish (*Betta splendens*) in a three-arm aquatic maze (Roitblat, Tham & Golub, 1982). The fish were placed in an arm chosen at random, and fed a single tubifex worm. Following this, they were placed in one of the other two randomly chosen arms. In the win-stay contingency, the original arm was rebaited and the fish was reinforced for returning to that arm. In the win-shift contingency, the arm that had not been previously entered on that trial (the third arm) was baited and the fish was rewarded for entering that arm. During the course of the experiment, each arm served in each role equally often. Half of the fish were trained in each condition for 70 trials (the maximum) or until they reached a criterion of eight correct responses (i.e. entries to baited arms) in 10 consecutive trials; following this they were trained with the other contingency. The fish were much faster in learning the win-shift contingency, requiring a median of 38 trials to reach criterion, than in the win-stay contingency, requiring a median of 70 trials to reach criterion.

Results such as these are usually interpreted as an indication of the organism's innate preference for one kind of behavioral strategy over another (e.g., a preference for relative novelty). They are also amenable, however, to the interpretation that the animals' access to, or formation of, representations about the location of food differed as a function of the contingency involved. Both the win-stay and the win-shift contingencies have the same formal requirements in terms of memory or other behavior; they contain the same amount of information and could be solved on the same basis. Yet one was easy for these fish and the other was difficult. This difference may be due to an inability to use the same information for two tasks. Similar results are reviewed below, indicating that rats in a radial arm maze also learn win-shift contingencies more readily than win-stay.

A similar phenomenon, involving an effect in the opposite direction, is the *feature positive effect* (Hearst 1978; Hearst & Jenkins 1974). The feature positive effect is seen in experiments in which the subjects are trained to discriminate two alternatives differing in a distinctive feature (e.g. the presence of a dot in a particular location on a pecking key). Discriminations in which the feature signals the positive alternative are learned more readily than discriminations in which the presence of the feature signals the negative alternative. This effect has been observed in a number of species, including rats (Halgren 1974), pigeons (Hearst & Jenkins 1974), and adult humans (Bitgood, Segrave & Jenkins 1976). Again, these results suggest that subjects differ in their ability to form or access representations of situations as a function of the contingency (task demands) involved.

Specification of representational systems

In this section I consider three task domains in which there has been considerable research into the representa-

tional system used. These tasks were chosen because they require an interaction between representations held in *reference* memory and representations held in *working* memory, and because the most analytical work has been done with these tasks.

Representation of a spatial task by rats

The first task to be considered involves spatial memory, as investigated in a radial arm maze. In this type of maze each of the arms extends from an elevated central platform, like the spokes of a wheel. Typically, a single piece of food is placed at the end of each arm at the start of a trial and the animal is allowed to run out along various arms. Once an arm has been visited, no more food is available there, so the optimal strategy is to choose each arm once per test session. In order to perform the task efficiently, the animal must have access to information about the arms already chosen, or about the arms where food remains.

Rats learn this task readily. On a maze with four arms they performed perfectly after about ten tests, choosing each arm once in the first four choices (Walker & Olton 1979). Rats also learned to perform the task in a 17-arm maze at well above the chance level, choosing (after about 50 tests) about 15 arms in the first 17 choices (Olton, Collison & Werz 1977).

There are a number of ways rats could solve such a spatial problem, some of which require the animal to maintain information about which arms do or do not contain food, others which only require the animal to maintain more global information about the task and to respond to the specific cues present as each choice is made. For example, if the rats were choosing different arms on the basis of odors emanating from the food, or on the basis of scent markers (e.g. urine) this would not be particularly interesting from a cognitive point of view. The odor, not a representation within the rat, would indicate the food's location. (The rat might have to represent the meaning of the smell internally and to process the sensory evidence internally, but the actual decision about which arm to enter would be made on the basis of external evidence immediately present.)

A number of experiments, however, indicate that these external cues play a negligible role. In one experiment (Olton & Samuelson 1976) rats were allowed to make three choices in an eight-arm maze and were then confined to the central platform while the arms were rotated and rebaited. Rats continued to choose arms that were in the appropriate spatial locations despite the fact that these same arms had previously been visited in a different spatial location.

Further evidence also indicates that the rats were representing the food locations relative to extramaze locations and that they had difficulty responding relative to intramaze locations when required to do so (Olton & Collison 1979). Rats in this study were trained in the standard manner to retrieve food from an eight-arm radial maze. Following acquisition to criterion, the rats were given one of two rotation treatments. For rats in one group, food was rotated with the maze, requiring the use of intramaze cues to keep track of it. For rats in a second group the food was placed on small platforms beyond the end of each arm; these arms did not rotate with the maze,

but instead stayed in constant locations relative to extramaze stimuli. Performance of the intramaze group never differed reliably from chance while the performance of the group for whom extramaze cues were relevant remained highly accurate.

Codes and contents. There are also data concerning the code and content of the representation used by rats in a radial arm maze. First, the possibility that the rats were using some kind of algorithm or response chain as the code was ruled out (Olton 1978). The difference between a response chain and an algorithm, as these terms are used here, is in the content of the representation. An algorithm is a relational code that specifies the next arm to be visited, based on its relationship to the current arm (the content). For example, if the first arm visited in a trial were called arm 0, then, independent of the physical location of arm 0, the relative location of arm 1 would be specified by the algorithm. In a response chain, on the other hand, the content of the representation would be the most recent arm entered and the code would specify where the next response should be made (e.g. as a stimulus response chain, in terms of right and left turns after each arm, or in terms of absolute intra- or extramaze physical locations). Using an algorithm would result in a particular pattern of choices relative to the starting arm. Using a response chain, on the other hand, would result in a particular pattern of choices relative to specific locations as well as the starting arm.

Neither of these alternatives has received much support. Although there was enough regularity in the pattern of choices to reject the hypothesis that the rats were simply sampling the various arms randomly without replacement (Olton 1978), the regularities were not sufficient to account for the high accuracy displayed. Furthermore, when rats were confined to the central platform for 20 seconds following each choice (Olton, Collison & Werz 1977), choice accuracy on a 17-arm radial maze remained high while the response pattern virtually disappeared. As further evidence, the probability of an error on choice 7 or choice 8 in an 8-arm radial maze was independent of the number of previous errors (Olton & Samuelson 1976; Olton 1978). The algorithm and the response chain coding hypotheses would both predict that errors would be more likely during the last two choices if there were more errors in the earlier choices, because each early error would effectively reset the algorithm or chain to continue from that particular arm on that choice.

Another possible candidate for a code used by rats in the radial arm maze is an analog of the arms in their spatial locations, that is, a cognitive map. Although this code has much intuitive appeal, it also has some difficulties. The simple presence of a map of the maze would not be sufficient for performance of the task. In addition to the map, which could be constant from day to day and therefore held in reference memory, task performance requires additional information about the arms still containing food, updated in working memory after each choice. This information could also be stored as a spatial analog, but the evidence available suggests that it is not. If the information were coded as a spatial analog, then errors would be likely to reflect a confusion of nearby spatial locations, and so would be expressed as entries to arms near the correct arm; but errors of this type occurred

no more often than would be expected by chance (Olton & Samuelson 1976). Instead erroneous choices were more likely to favor arms that were visited early in the trial than arms visited later (Olton 1978; Olton, Collison & Werz 1977; Olton & Samuelson 1976; Roberts & Smythe 1979). If the information were encoded in terms of arms still containing food, then errors would be expected to be distributed randomly, relative to serial position, because the information concerning which arms had been visited early, as opposed to late, would not be available. For serial position to affect error distributions, serial position information would have to be part of the representation. These results suggest that the information is coded in working memory as a list of previously visited arms, rather than as a direct analog of spatial location.

If some kind of map of the physical location is involved in rat radial arm maze performance (e.g., in reference memory), it must, if it is to be useful, be augmented by some kind of compass specifying the orientation of the map relative to the environment. One likely candidate for such a compass is the complex of extramaze stimuli typically present in the experimental room (e.g. windows, sinks, etc.). In order to investigate the role played by these stimuli, Suzuki, Augerinos and Black (1980) carefully enclosed an eight-arm maze inside a large cylindrical chamber. In their first experiment the choice accuracy of rats was compared both when distinctive stimuli were hung inside the cylinder and when no distinctive stimuli were present. Throughout training the group with the distinctive stimuli made fewer errors than did the group trained without distinctive stimuli, but no statistics are presented to indicate whether this difference was significant.

In their second experiment each trial consisted of three stages: (1) three forced choices to arms selected by the experimenter (access was controlled by means of guillotine doors; different arms were chosen on each trial); (2) confinement to the central platform during stimulus manipulation (2.5 min); (3) free access to all arms. Before each trial the spatial relationships among the cylinder, the stimuli, the arms and the central platform were varied to prevent the animals from using stimuli other than the experimental stimuli.

One of three stimulus manipulations was used during each trial. On *control trials* the stimuli were rotated around the maze and returned to their original locations. On *rotation trials* the stimuli were rotated along with the walls and ceiling of the cylinder 180° and left there. On *transposition trials*, the stimuli were exchanged. After rotation or transposition the arms associated with previously unchosen stimuli were baited. A correct choice was counted as a choice to an arm associated with a previously unchosen stimulus (during the forced and/or the free choices), independent of its location.

Choice accuracy on the rotation trials did not differ from choice accuracy on control trials; the rats tracked the rotation, continuing to choose arms associated with previously unchosen stimuli. Choice accuracy following transposition, however, declined to a level near chance. "Analyses of errors suggest that in the transposition trial the animals did not avoid the stimuli chosen in the forced choices and performed as if a new trial had been started after stimulus transposition." (Suzuki et al. 1980, p. 11)

Furthermore, in their third experiment Suzuki et al. found that the stimulus manipulation did not affect the number of different arms chosen in *eight* choices after confinement, and that the choices made after confinement were independent of those made before confinement during transposition, but not during control or rotation trials.

Their tracking of the stimuli following rotation clearly indicates that the rats were depending on the experimental stimuli as the compass for orienting themselves relative to the maze. Suzuki et al. (1980) argue on the basis of this finding that the rats treated the stimuli as a configuration, and that the failure to track in the same way following transposition indicates that the rats did not identify each arm by its associated stimulus, but by its relation to the configuration. These results, they claim, are incompatible with a working memory representation based on a list because, in their view, a list requires that each arm be identified by a single extramaze stimulus. There are two problems however, with this conclusion. First, as Suzuki et al. recognize, nothing in the list hypothesis requires that a location be identified by a *single* stimulus; any combination would have been disrupted by the transposition treatment, but not by the rotation treatment. Second, the rats treated transposition as a signal for the start of a new trial, carrying over no information from the forced choices. Whatever other function the stimuli were serving may have been masked by this resetting function. These results, and the previously described results, are compatible with the hypothesis that the working memory representation consists of a list of spatial locations. As each choice is made, another spatial location is added to the list. Spatial locations, in turn, are maintained in reference memory in a representation that has properties similar to a map, and is oriented relative to the complex of extramaze stimuli.

Dynamic properties. The most striking feature of the dynamic properties of rats' representational system for radial arm maze performance is its durability over long delays. Rats were able to complete an eight-arm maze four hours after completion of the first half (i.e. four choices), with better than 90% accuracy, and their performance remained above chance with delays as long as 12 hours (Beatty & Shavalia 1980).

When a large number of treatments, known to affect performance on other tasks, were presented immediately (i.e. within 2 minutes) after a fourth choice (i.e. in the interval between the fourth and fifth choice), they had surprisingly little effect on the remaining four choices. Neither variations in illumination or auditory levels, removal from the experimental room, feeding, or exposure to a four-arm maze were effective in reducing choice accuracy on the last four choices (Maki, Brokofsky & Berg 1979).

Exposure to a second 8-arm maze during the retention interval between choices 4 and 5 also had no effect (Beatty & Shavalia 1980). In the first experiment, eight choices on a second maze intervened between the fourth and fifth choices on the first maze. Beatty and Shavalia reasoned, however, that this treatment's ineffectiveness may have resulted from allowing the rats to complete the second maze before returning to the first. There was thus no reason for the rat to retain information about the second

maze when returned to the original. To test this idea, the rats were allowed four choices in maze A; then, after a varying delay, they were allowed four choices in maze B. Four hours after their fourth choice in each maze they were returned to that maze and were allowed four more choices. Choice accuracy in both mazes was uniformly high and not significantly different from performance on those control days that involved only one maze and a four-hour delay.

Administration of electroconvulsive shock (ECS), however, has been found to affect radial arm maze performance (Shavalia, Dodge & Beatty 1981). Rats in this series of experiments were administered ECS at varying times during a four or six hour retention interval between the fourth and fifth choices. Administration within 15 minutes following the fourth choice had no effect on the subsequent choices. If the treatment was delayed for 30 minutes, then some decrease of choice accuracy on choices five through eight was obtained. Severe amnesia was produced by presentation of ECS two or four hours after the fourth choice.

The errors obtained following effective ECS treatments consisted of reentries into arms visited on choices one through four but not into those arms entered after the delay. This finding, coupled with the results obtained from control tests in which ECS was administered two hours prior to the first four choices, indicates that ECS had its effect by retroactively interfering with the representation of the first four choices. The increase in the disruptive effects of ECS with increasing delay suggests two possibilities: First, ECS treatments might function to decrease memory strength (clarity, etc.) by a fixed amount independent of current strength. Early administration of ECS acts on a very strong representation leaving enough strength to support high levels of performance after the delay. After two hours, however, the current memory strength is low enough so that a decrement of similar magnitude practically eliminates the memory trace. Alternatively, the information about the first four choices might be maintained for a short time after the first four choices in a form that is resistant to ECS interference. At some point during the delay, it may be translated from that ECS resistant form to some other form which might be more resistant to normal forgetting but less resistant to ECS. This translation process, if it exists, appears to have a significant effect some time beyond 15 minutes past the fourth choice. If the partial interference of ECS obtained after a 30-minute delay is obtained with individual rats, and is not the result of averaging, then it suggests that the translation process might be gradual.

At this point, there are not sufficient data to decide between these alternative models for the effect of ECS. If ECS were to reduce the strength of a memory by a constant amount, independent of time of administration, then one would expect the performance to reflect that decrement in lower levels of accuracy. In this experiment, however, early administration of ECS had no effect – neither one favoring the translation over the constant decrement hypothesis, nor one suggesting that the small decline was insufficient to be seen in behavior. Whatever the ultimate outcome, the interval from 15 to 120 minutes after the fourth choice appears to be very important.

Counterpoint to the apparent durability of the repre-

sentation is the finding that the memory system can apparently be reset (see also Suzuki et al. 1980), thereby attenuating proactive interference from earlier choices on the same maze. Rats were tested for eight consecutive trials, each separated from the next by one minute (Olton 1978). Although choice accuracy declined as the number of choices made within a trial increased, it returned to errorless levels at the start of each trial.

A closer examination of Olton's data, however, revealed that errors began appearing earlier in the second trial of a test than in the first (Roberts & Dale 1981), thus indicating the presence of some proactive interference on the second trial. Roberts and Dale replicated Olton's finding but found that no further changes were apparent after the second trial. They also found that repeated experience with multiple trials per day attenuated the disruptive effect but increased the likelihood that the rat would choose adjacent arms over the course of a five-trial session and over the several days, thereby suggesting that the rats adapted to massed trials by adopting an algorithmic strategy.

These data and the data presented by Beatty and Shavalia (1980) suggest that rats can discriminate one trial from another and, to some extent, keep separate the information from the two. Because performance remains at a high level over long delays, loss of information from an earlier trial cannot be attributed to the passage of time, but is more probably due to some imperfect, directed forgetting mechanism that depends on the ability of the animals to discriminate successive trials from one another (see also Suzuki et al. 1980).

Summary. The following is proposed as a summary of the data presented above and as a tentative description of the representational system used by rats in the domain of radial arm maze performance. The content of the working memory representation appears to involve the location of previously visited arms relative to extramaze environmental characteristics. The medium appears to be neural/cognitive. The content of memory consists of a map of spatial locations maintained in reference memory and a list of visited locations maintained in working memory. The role of external stimuli is to orient the map, not to serve as specific cues for specific arms. Although the code cannot be confidently specified, serial position effects on the likelihood of reentry suggest that it has temporal components related to the order in which arms were chosen. The effects of ECS suggest (perhaps) that the particular code used may change during the course of a delay. Although the coded representation is ultimately susceptible to ECS-produced interference, it is otherwise remarkably durable.

Representation in delayed matching-to-sample

Delayed matching-to-sample (DMTS) has been the most widely used task for studying animal memory. A number of different species have been tested in this task including monkeys (D'Amato & Cox 1976; Mishkin & Delacour 1975) and dolphins (Herman & Gordon 1974). The amount of research done on DMTS is far too vast to be covered here, so the following discussion will be limited mainly to the more analytic work involving pigeons. DMTS involves the presentation of a sample stimulus,

followed by a retention interval and the simultaneous presentation of two or more comparison, or test, stimuli. Responses to the comparison stimulus that matches the sample stimulus are rewarded. Because the sample stimulus is no longer present when the comparison stimuli are presented, accurate performance relies on the perseveration of the information contained in the sample until a correct choice can be generated.

Codes and contents. As the name of the task implies, DMTS is most easily described in terms of a matching concept: Choose the comparison stimulus that *matches* the sample. For this reason theoretical interpretations of the performance have also tended to describe the process in similar terms. Thus, the increase in performance found with increasing sample duration (Maki & Leith 1973; Roberts 1972; Roberts & Grant 1974) and the decline in performance typically found with increasing delay, have both been ascribed to an increase and subsequent decline in the strength of a unitary memory trace, assumed to be a more or less direct copy of the sample stimulus. When the comparison stimuli are presented, "a subject compares the stimuli presented on the side keys (the comparison stimuli) with a trace of the sample stimulus stored in memory. If a match between trace and test stimulus is made, a correct choice results." (Roberts & Grant 1976, p. 88). This view assumes that sample presentations are coded in working memory as a more or less direct copy of the sample stimulus, which grows in strength with increasing sample duration, and declines with increasing delay between sample and comparison stimulus presentations. Matching is assumed to depend on the physical relation (similarity) between the comparison stimuli and the remaining trace of the sample stimulus.

Three lines of evidence converge on the conclusion that this copy model is too simple to account for all DMTS performance. First, physical similarity between sample and comparison stimuli is not sufficient for accurate matching performance (Cumming, Berryman & Cohen 1965). Pigeons were trained in a DMTS task in which the correct test stimuli were identical to the sample stimuli, but were either novel or familiar. If physical similarity between sample and test stimuli were sufficient to produce high levels of matching accuracy, then matching accuracy with a novel, sample-test pair would have been as good as with previously trained pairs. However, choice accuracy with a novel sample was initially at or below chance. Apparently some experience with the pair is necessary before accurate responses can result (cf. Riley, Cook & Lamb 1981).

One role that prior experience may play in determining DMTS performance is in the formation of attentional or coding strategies for the relevant stimuli. This hypothesis receives some support from an experiment reported by Maki, Gillund, Hauge and Siders (1977). Pigeons learned a two-choice, 0-delay matching-to-sample task in which the samples were either colors or line orientations, but only colors were used as comparison stimuli. Following color samples, one test stimulus exactly matched the sample and was designated the correct choice; the other test stimulus was the second color. Following line orientation samples, one color (e.g. red) was arbitrarily designated correct following each (e.g. horizontal) sample (See Table 1). The latter procedure is called *symbolic match-*

Table 1. *Stimuli used by Maki, Gillund, Hauge & Siders (1977)*

Sample	Correct choice	Incorrect choice
red	red	green
green	green	red
vertical	red	green
horizontal	green	red

ing because the correspondence between the sample and test stimuli is arbitrary rather than based on similarity.

Pigeons were normally required to make ten observing-response pecks to the sample stimulus in order to initiate the presentation of the test stimuli. Following initial training, however, one of the sample stimuli was arbitrarily designated as the to-be-extinguished sample and was presented for a fixed duration of 5 seconds, during which pecks had no programmed consequence. Observing responses to this sample stimulus no longer led to presentation of the choice stimuli and quickly dropped out; however, observing responses to the other (still tested) samples continued. Finally, a few probe trials were presented in which the extinguished sample was followed by test stimuli. Choice accuracy on these probe trials was at the chance level, indicating that at the time of testing, the bird had no information regarding the identity of the sample stimulus.

This finding is remarkable for a number of reasons. First, on half of the trials the sample stimuli were line orientations but the test stimuli were colors. The high level of performance obtained during initial training (as well as that obtained in other symbolic DMTS studies) indicates that a physical match is not necessary for matching performance.

Second, the chance performance obtained following the extinguished sample was not due to a simple avoidance of the associated test item. Because of the mixture of identity and symbolic DMTS, the correct choice item was also associated with another sample stimulus on which choice accuracy remained high.

Third, these results indicate that the bird had identified the sample well enough to avoid pecking at it, but had failed to perform whatever operations (e.g. encoding, rehearsal, etc.) were necessary to produce a correct matching response. Something more than simple identification is necessary.

The copy theory (Roberts 1972; Roberts & Grant 1976) suggests that pigeons perform DMTS on the basis of some kind of matching concept, for example, a rule of the form "pick the test item most similar to the sample stimulus whose copy appears in memory." Although there is evidence that pigeons are capable of using a general rule concept (Herrnstein, Loveland & Cable 1976; Wright, Santiago, Urcuoli & Sands 1981; Zentall & Hogan 1978; Zentall, Edwards, Moore & Hogan 1981), its use appears to require special training and is not necessary to DMTS performance. More typically, and necessarily with symbolic DMTS, pigeons appear to base their choices on particular sample-comparison stimulus pairings. According to this multiple rule model, the subject maintains in reference memory one rule for each sample-test pair

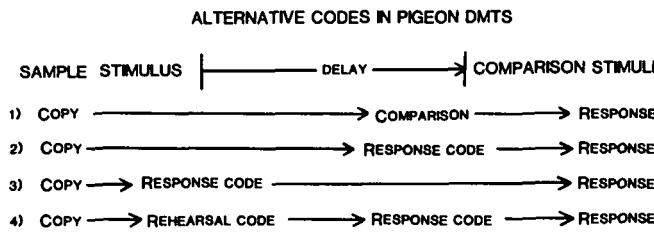


Figure 2. Alternative memory codes pigeons could use in symbolic delayed matching-to-sample.

(Carter 1971; Carter & Eckerman 1975; Carter & Werner 1978; Cumming & Berryman 1965). The rule specifies the mapping of comparison stimuli onto sample stimuli, that is, the correct choice. On every trial the bird must remember which sample stimulus occurred on that trial. Through the application of the rule to that information, the correct choice can be determined. This translation process from sample to correct choice must occur some time between the onset of the sample stimulus and the time a choice is made (See Figure 2).

The information could be maintained in any of at least three forms (Roitblat 1980):

1. Sample code. The bird could remember a more or less direct copy of the sample stimulus during the delay and apply the rule to it only when the choice stimuli were presented. This is a restatement of the simple copy theory (Roberts 1972), with the additional assumption that choices are made on the basis of a mapping rule rather than on the basis of a direct comparison of the sample and test stimuli. According to this alternative, the mapping rule is applied late in the interval between sample and test, just prior to the occurrence of a choice. For most of the delay, then, memory is coded as a copy of the sample stimulus (Code 2, Fig. 2).

2. Test code. The bird could immediately translate the sample stimulus into a form isomorphic with the correct test stimulus and maintain that representation in memory during the delay. The initial copy of the sample stimulus is quickly translated (recoded) into a form resembling a copy of the correct comparison stimulus, in which form it is maintained for the rest of the delay and which eventually provides the basis for responding (Code 3).

3. Intermediate code. Finally, the information could be maintained in some intermediate form intervening between the sample and the test in the translation process. The initial copy of the sample stimulus is quickly translated according to some mapping rule into an intermediate code resembling neither the sample nor the correct comparison stimulus, maintained in that form for most of the delay, and finally used as the input to a second mapping rule that produces the choice. In humans such an intervening form might be a verbal label, for example. In animals, this intermediate form might be some kind of "time tag" (D'Amato 1973; Mason & Wilson 1974; Mishkin & Delacour 1975; Winograd 1971; Worsham 1975) described more fully below (Code 4).

The particular technique used to discriminate among these three alternative memory codes is based on an experiment by Conrad (1964) in which people were asked

to remember visually presented consonants. The errors that occurred tended to be acoustically similar, not visually, to the correct item. Based on the assumption that confusions should be more likely between similar than dissimilar items (Beals, Krantz & Tversky 1968), Conrad inferred that the representational code used by humans in his task bore more of a similarity to the acoustic than to the visual properties of the memory items.

A technique based on similar logic was employed to investigate the code used by pigeons in a symbolic DMTS task (Roitblat 1980). Color sample stimuli were followed by line orientation test stimuli, and, for a different bird, line orientation sample stimuli were followed by color comparison stimuli. Three samples and comparison stimuli were used; one sample and three comparison stimuli appearing on each trial. The sample and comparison stimuli were chosen so that similar comparison stimuli corresponded to dissimilar sample stimuli, and dissimilar comparison stimuli corresponded to similar sample stimuli.

For two of the birds, sample B (orange) was dissimilar to sample A (blue) but similar to sample C (red). The comparison stimuli had a complementary similarity relationship. Comparison item B (12.5°), the correct choice following sample B, was similar to comparison item A (0° , vertical) and dissimilar to comparison item C (90°). For the third bird the samples and comparison stimuli were interchanged.

For the two birds trained with color samples and line orientation tests, an increasing retention interval resulted in an increase in the number of errors. The pattern of those errors reflected a faster increase in the rate of confusion between similar comparison stimuli corresponding to dissimilar samples than between dissimilar comparison stimuli corresponding to similar samples (See Figure 3). The third bird, trained with line orientation samples and color comparisons, showed a similar trend, but the change was not significant. The change in confusion with an increase in the retention interval is most consonant with the test code hypothesis presented above. Finding only a higher rate of confusion between similar, than between dissimilar, comparison stimuli might reflect only a difference in the discriminability of the stimuli. The change in confusion due to delay, however, cannot be similarly interpreted since the stimuli themselves did not change during the delay; rather something in the birds' memory did. As the memory was maintained longer, more opportunities for confusing representations were available.

The working memory representational system used by pigeons in the domain of DMTS has the following features: (1) The medium is usually internal to the animal, although external, or behavioral, codes have been observed and trained. In the confusion experiment (Roitblat 1980) just described, the evidence strongly favors an internal code because, to produce the observed pattern of confusion errors, any external mediator would have had to have properties similar to line orientation for two birds, and to color for a third. (2) The content of the representation appears to be the stimulus to which the animal should respond during the test phase, coded in a form isomorphic with the dimension of variation of the comparison stimuli.

The representational system suggested above seems to

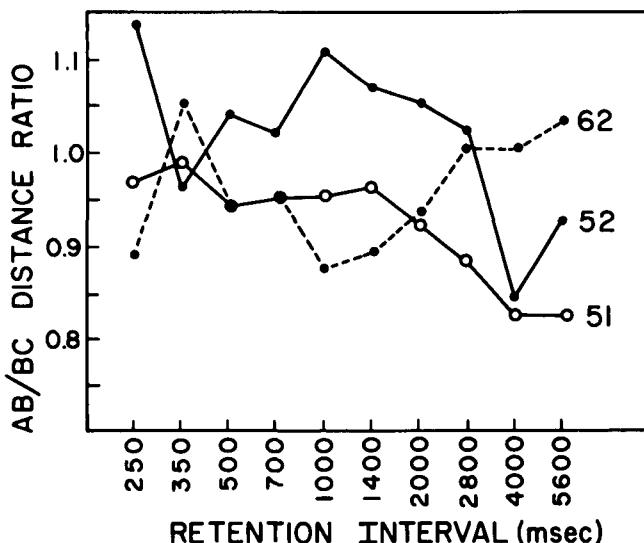


Figure 3. The ratio of the discriminative distances between stimuli A and B relative to the discriminative distance between stimuli B and C. Discriminative distances are assigned in a manner analogous to that used in multidimensional scaling. Distance declines as confusion between the two stimuli increases. For birds 51 & 62 samples were blue, orange, and red (A, B, C, respectively), corresponding respectively to comparison stimuli 0°, 12.5°, and 90° line orientations. For bird 62 the sample and comparison sets were exchanged relative to the other two birds so, from a theoretical point of view, an equivalent change is in the opposite direction for this bird.

be different from that employed by monkeys. D'Amato and others (D'Amato 1973; D'Amato & Cox 1976; Mason & Wilson 1974; Worsham 1975) have suggested a working memory representational system in which the information about sample presentations is coded in working memory as the relative recency of each sample's occurrence (time tag). This view is motivated by the following: performance by monkeys on DMTS improves with increasing intertrial intervals (Jarrard & Moise 1971) and with increases in the size of the sample stimulus pool (Mason & Wilson 1974; Worsham 1975); it declines with increases in the frequency with which the stimuli appear within a session (Mason & Wilson 1974). Performance on "hard" sequences of trials, in which the sample from one trial occurs as the incorrect comparison stimulus on the next, is poorer than performance on "easy" trials, in which the incorrect comparison stimulus had not recently been seen (Worsham 1975). Finally, following extensive experience, the choice accuracy shown by monkeys appears to be independent of sample duration (D'Amato & Worsham 1972).

The time tag hypothesis relies on a working memory code that is an analog of the relative recency with which each of the possible sample stimuli occurred, instead of an analog of the stimuli themselves. This hypothesis depends on the stimuli having been well learned (i.e. sufficiently coded into reference memory). Once the stimuli are familiar, "each stimulus could become a highly organized configuration which could be identified even with partial information" (D'Amato 1973, p. 246). Sample duration is thought to play a role only to the extent that the stimuli are unfamiliar.

Although relative recency plays a role in the DMTS performance of monkeys, the evidence for an analog of

recency as the memory code is not conclusive. All of the above findings are also consistent with the formation of a representation that varies in strength, increasing with sample duration and declining with retention interval duration. If the representations from the last several trials compete for control of the choice, then performance would improve with increases in the intertrial interval, and with increases in the size of the sample pool, because the representations would differ more both in strength and in relative recency. Similar arguments apply to "difficult" versus "easy" trial sequences. Finally, with extensive experience the monkeys might be learning to limit the number of features of the sample stimulus information that are encoded, or to encode them more efficiently, hence forming adequate representations more quickly. In that case, it is not that sample duration has no effect, but rather that the encoding process becomes so fast that it reaches a ceiling in less than .075 second, the briefest duration tested.

In order to decide the issue, features of performance other than the average number of errors should be examined. For example, if the pattern of confusion errors depends systematically on features of the sample and comparison stimuli, even after they are well-learned, then those features are likely to have been represented in working memory.

Sequential representations

A third area in which the representation used by animals has been explicitly studied involves memory for sequences of events. The traditional means of studying sequential processes in animals involved spatial maze problems such as multiple T-mazes. In tasks such as these, however, there is no reason to assume that the subjects are using representations of the order of events, or of responses. A simple unordered representation of choice-point/response combinations would be sufficient (Bever, Straub, Terrace & Townsend 1980). Other approaches are necessary, therefore, to explore how animals form representations of ordered sequences.

Representation of patterns of reinforcer magnitudes. Hulse and his associates have been studying the interaction between the order of presentation and the alphabetic (i.e. pre-experimental) ordering of sequences consisting of reinforcer magnitudes. In an initial experiment with rats (Hulse & Campbell 1975) the pattern consisted of various orderings of 0, 1, 3, 7, and 14 food pellets at the end of one arm of a T-maze. The first run of each trial consisted of a brightness discrimination. Choice of one stimulus (S^+) was rewarded with the first element in the series. For the second and subsequent runs the unchosen arm was blocked off and the next element in the series was presented on each subsequent run as a forced choice. If the other stimulus (S^-) was chosen on the initial free choice, then one more unrewarded forced choice of that arm was given, followed by the intertrial interval.

Hulse and Campbell (1975) compared the running latencies (i.e. the time needed to get from the start to the goal box) to each element in the series as a function of different patterns of elements. One group received the elements in ascending order: 0 pellets on the first free

choice, 1 pellet on the first forced choice, 3 pellets on the second forced choice, etc. Another group received the same elements in descending order, with three more groups receiving random orders ending in 0 pellets for one, 14 pellets for a second and a random number for the third. Rats in the increasing and decreasing order groups decreased their latencies as a function of reward magnitude: The more pellets in the element, the faster they ran. The random groups, on the other hand, showed no evidence of sensitivity to reward magnitude.

A number of different representational systems have been suggested to account for these findings. Neither hunger, temporal associations, nor counting provide an adequate basis for performance (Hulse 1978). Two other systems remain viable. Hulse has proposed a "pattern rule" representational system in which the rat is assumed to use features of the sequence's formal properties, such as monotonicity, to encode the series as a pattern or rule (Hulse 1978). In contrast, Capaldi and his associates (Capaldi, Verry & Davidson 1980a) have argued for a representational system in which the sequence is encoded as a set of stimulus-stimulus associations. The memory for each element in a series is assumed to serve as a cue signalling the next element. The signalling capacity of a stimulus depends on its association with the subsequent element, and on generalized strength obtained as a function of its similarity to other elements in the sequence.

Both models propose that the sequence in which the stimuli are presented interacts with the ordering of the stimulus magnitudes. Both require that the animals order these magnitudes. For the association/generalization model the interaction takes the form of generalization from each stimulus to the others. For the pattern/rule model the interaction consists of an extraction of simple or complex rules. In addition, Capaldi's model requires that the rats represent in reference memory information about the interval level differences among the stimuli. Although Capaldi's association/generalization model appears to be simpler than Hulse's, relying on paired associates rather than on some abstract rule, it gains its power not from the association mechanism but from generalization, and as a result actually requires the animal to maintain more information about the stimuli than is required by a rule-based model, which requires only that the stimuli be ordered. The association/generalization model, on the other hand, requires that they be scaled in order to utilize similarities between items as the basis for generalization.

The available data favor the ordinal-rule-based interpretation over the association/generalization interpretation, although the critical experiments have not been run. (Despite the authors' claims to the contrary, the experiments presented by Capaldi et al. 1980a, for example, involve a different task from Hulse's, making it difficult to interpret them. Until it can be shown that their procedure can produce findings analogous to those obtained with Hulse's procedure, any assertions are simply groundless.)

Strongly monotonic patterns (14-7-3-1-0) were learned more quickly than weakly monotonic patterns (14-5-5-1-0) (Hulse 1978; Hulse & Dorsky 1977) and more quickly than nonmonotonic patterns (monotonic: 18-10-6-3-1 or 18-10-6-3-0; versus nonmonotonic:

18-3-6-10-1 or 18-3-6-10-0; Hulse & Dorsky 1979). The difference between the strong and weak monotonic conditions might be partially explained by a pair-wise theory that suggests that the repetition of the 5-pellet element results in an association between it and both another 5-pellet element and a 1-pellet element. The resulting confusion is not sufficient, however, to account for the slower differentiation of running speed to the last 0-pellet element. In both conditions, this last element was preceded solely by a 1-pellet element, so that once the 1-pellet element was experienced, no confusion would be expected. Furthermore, Capaldi's model (as I understand it, see below) predicts better differentiation between running speeds to the 0-pellet element and to the others in the weakly monotonic case because the 1-pellet element gains less generalized signalling capacity when the two 5-pellet elements are included in the series than when they are replaced by a 7- and a 3-pellet element. Finally, all repetition was eliminated in the nonmonotonic condition, but learning was again retarded relative to a monotonic pattern of identical elements.

Other evidence indicates that simpler formal properties not only make the sequence easier to learn but are also maintained in the representation. Rats were transferred from either a monotonic or a random sequence of reinforcers to either a monotonic or nonmonotonic sequence (Hulse & Dorsky 1979). During training the monotonic group received sequences of either two, three or four elements, drawn from the set 0, 1, 3, 5, or 10 pellets. All sequences for this group were monotonically decreasing, with the restriction that the subsequence 1-0 was never presented. The random group was yoked to the monotonic group in the sense that the same elements in identically-sized lists were presented to both groups, but the order of the elements was randomized. After this training for 70 trials, half the animals in each group were transferred to a new monotonic pattern (16-9-3-1-0) or to a new nonmonotonic but stable pattern (16-1-3-9-0).

Learning of the second pattern was measured by the number of trials necessary to detect a reliable differentiation of the 0-pellet element and by the ultimate (after 25 trials) increase in running time for the last elements in the series. Learning of the transfer task was fastest for the monotonic-monotonic group, relative to either of the randomly pretrained groups, and slowest for the monotonic-nonmonotonic group, indicating transfer from monotonic pretraining. These results suggest that rats maintained information about the formal structure of the sequence that affected later performance.

To the extent that the association/generalization model fits these data, it relies on generalization from each element in the sequence to the others. Because the generalization function has not been explicitly presented, it is difficult to assess the adequacy of Capaldi's interpretation. I have made simple assumptions, derived from Blough's (1975) analysis of generalization mechanisms, about the generalization functions and about reward values (utilities); I have used these assumptions to generate predictions for two of the groups in Hulse and Dorsky's (1979) experiment. The first group was trained with the sequence 18-10-6-3-0 and the second with the sequence 18-3-6-10-0. Capaldi et al. (1980a) claim that the association/generalization model, like the rule-based model, predicts that the nonmonotonic pattern should be

Table 2. Predictions of the generalization/association model

A. Amount of generalization		To		
From		18	10	6
18	1	.18	.35	.13
10			1	.35
6				1
3				

B. Scaled reward utility				
Pellets	0	3	6	10
Utility	0	1.1	1.8	2.3

C. Generalized expectancy value V(I) for each element in the sequence				
Element Presented	Contribution from element			V(I)
	18	10	6	3
18	2.3	.63	.14	0
10	.80	1.8	.38	0
6	.30	.63	1.1	0
3	.04	.23	.38	0
	18	3	6	10
18	1.10	.03	.30	0
3	.01	1.80	.80	0
6	.14	.63	2.30	0
10	.38	.23	.80	0
				1.41

Note: Generalization factors were extracted from Blough (1975). The model makes no explicit prediction about speed to the first (18-pellet) element, but predicts running speed on the second through fifth elements as a function of the previous element. Running speed on the run following the I-pellet element is a function of V(I).

more difficult to learn and, at asymptote, should result in a smaller difference in running speed to the terminal 0-pellet element than with the monotonic pattern. This prediction (based on my assumptions) is presented in Table 2. No explicit prediction is made about running speed to the first element in the series, because the amount of generalization from the cues signalling the start of the run to the other stimuli in the sequence is unclear. Once the first element has been obtained, however, running speed should be proportional to the values labelled V(I) in the table for each successive run. (Speed must be proportional to, rather than simply monotone with, V(I) in order to predict a difference between monotonic and nonmonotonic patterns, since ordinal relations are preserved in both conditions.) As can be seen in Table 2, running speed in the nonmonotonic case is predicted to be low following the 18-pellet element, then rising until the 10-pellet element is received, after which it is predicted to drop to a level only slightly below that following the 18-pellet element. The model thus predicts less differentiation between the 3-pellet and the 0-pellet element, but more differentiation between the 0-pellet and the preceding element in the nonmonotonic condition. This pattern is not consistent with the data presented by Hulse and Dorsky (1979), and is still predicted if somewhat stronger or weaker generalization is assumed. (Too little generalization results in the prediction that

monotonic and nonmonotonic patterns should not differ and too much generalization results in other predictions that are not consistent with the data.)

Sequence production. A recent series of experiments, involving another approach to sequential processing, demonstrates that pigeons can produce a sequence of responses to a simultaneously presented set of four colors (Straub 1979; Straub, Seidenberg, Bever & Terrace 1979; Straub & Terrace 1981; Terrace 1980), independent of their physical location. All stimuli were presented simultaneously and no explicit feedback was given except following an error (pecking one of the colors out of sequence), which terminated the trial, or following a correct completion of the sequence, which produced access to food. With extensive training the birds learned to choose the four colors in order at a level of reliability well above chance.

There are at least four ways the birds could have been representing this sequential behavior. According to the configural hypothesis, the birds learned each sequence as specific stimulus-response chains in which the stimuli were spatially ordered four-color configurations and the responses were complete sequences. To test the configural hypothesis, Straub et al. trained their birds with a subset of 15 of the 24 available configurations in a 4-key box, and then tested with four of the remaining nine novel configurations. There was very little decrement relative to familiar arrays (replicated by Straub 1979), thus ruling out the configural hypothesis.

According to the second hypothesis, the sequence was coded as a chain of stimulus-response-stimulus-response units such that the response to one stimulus produced characteristic feedback, which was then associated with the next response. According to a similar, third hypothesis, the bird coded the sequence as a series of unordered conditional rules. "Peck stimulus A, having just pecked stimulus A, peck stimulus B, etc." (Straub et al. 1979).

Data forcing rejection of both chaining hypotheses was obtained through the use of subsequences (Straub 1979; Straub & Terrace 1981). Following initial training with the entire array of stimuli, each bird was tested with subsequences of the pattern (e.g. A-B, A-D, B-D, A-B-C, A-B-D, B-C-D, etc.). Both chaining hypotheses predict that performance on subsequences involving nonadjacent stimuli, or arrays that do not start with the first element, would be very poor. Performance, however, was good on all subsequences tested.

At this point, the most reasonable conclusion is that the birds represented the entire sequence as an ordered set. The data presented by Straub (1979) and Straub et al. (1979) are not consistent with the simple associative chaining models described above, but they are well predicted by a simple production model (Bever et al. 1980) based on the assumption that an ordered representation of the stimuli in sequence (content) is scanned by the subject during production of the sequence.

Serial probe recognition. There is ample evidence that animals can represent sequences of presented stimuli in working memory. For example, a monkey was trained in a serial probe recognition (SPR) task to make one response if a probe item was presented earlier in a list, and another if the probe item had not occurred in that session

(Sands & Wright 1980). The lists were drawn from photographic slides of a wide variety of items such as fruits, flowers, people, etc., and were changed after every four sessions. The monkey was quite accurate in this task, showing both a primacy and a recency effect: Accuracy with probe items corresponding to stimuli presented in the middle of the list was lower than accuracy with items presented either early or late in the list. A similar serial position effect was obtained from a human in the same task.

Serial probe reproduction. In another experiment, with pigeons, stimulus order not only affected the representation, but was also found to be contained in the representation. In this experiment a series of stimuli was presented followed by a probe for one of those stimuli cued by its serial position. Shimp (1976b) trained pigeons to peck at a three-item sequence of X-shaped forms, each presented for a fixed duration in random order on two side keys of a three-key Skinner box. After a retention interval, the birds were cued to report the stimulus in a particular serial position by first illuminating the center key with a particular color, then, following a peck to that key, by illuminating both side keys with the same color. Red, for example, indicated that pecks to the side key on which the X had first occurred would be reinforced. Blue and white indicated the second and third serial position respectively. Choice accuracy was well above chance for most presentation durations, and most retention interval durations. As was found in the monkey SPR experiment (Sands & Wright 1980), a strong serial position effect of recency was apparent when all three stimuli were presented for .5 seconds separated by a .1 second ISI, but no primacy effect was seen unless the duration of the first item was increased to 4 or 10 seconds while the remaining two stimuli were presented for .5 second each. These results strongly suggest that pigeons can represent in working memory the individual items and their sequence.

In a similar experiment, two monkeys were trained to reproduce a list of two presented stimuli through the delayed matching-to-successive-samples (DMTSS) technique (Devine, Burke & Rohack 1979). Two samples (colors or shapes drawn from a pool of four each) were presented to the monkeys, each for 2 seconds separated by a 0.3 second interstimulus interval (ISI). Following a delay of 4, 8, 16, or 32 seconds, three comparison stimuli were presented: one matching each of the two samples, and a third distractor stimulus that had not previously appeared on that trial. The monkeys were reinforced for responding to the two matching comparison stimuli in the order in which they had occurred as samples. Incorrect choices terminated the trial.

Following extensive training, the monkeys performed this task with better than 80% accuracy at minimal delays. They had great difficulty, however, with those sequences in which the two sample stimuli and the distractor were both drawn from the same dimension (e.g. color-color as sample and a third color as distractor). Performance on these same-dimension sequences was dropped from consideration for one of the animals and was considered only during a second phase of testing for the other.

For one of the monkeys, choice accuracy following all combinations of stimuli declined across delays. For the

other monkey, choice accuracy on those sequences consisting of dissimilar samples (one color and one shape) remained at 90% and better after all retention intervals, but following sequences consisting of similar samples it declined with delay. For both subjects, 90% of first choice errors were to the stimulus corresponding to the second sample rather than to the third, distractor stimulus. Furthermore, these errors were far more prevalent when the two samples were drawn from the same dimension than when they were drawn from different dimensions. Thus, these errors primarily reflected a loss of order rather than of identity information, suggesting that both item and order information were maintained, but stored independently.

These results indicate that pigeons and monkeys can represent both the item and the order information present in a serially-presented list. Little is known, however, about the code used by subjects in these SPR tasks, except that item and order information appear to be represented independently of one another (at least by monkeys responding to sequences of well-learned stimuli).

Sequence discrimination. Another approach to string processing in animals involves string recognition or delayed sequence discrimination (Wasserman, Nelson & Larew 1980; Weisman & Dodd 1979; Weisman, Wasserman, Dodd & Larew 1980). In this task a sequence of stimuli, separated by an interstimulus interval (ISI) is presented, and followed, after a retention interval (RI), by a test stimulus. Response to the test stimulus is reinforced following positive sequences, but not following other sequences. In one set of experiments with pigeons (Weisman & Dodd 1979; Weisman, Wasserman, Dodd & Larew 1980), three stimulus events were possible: two colors, "A," and "B," and an unfilled interval of similar duration, "X." Stimuli were presented either on pecking keys or through colored houselights. The sequence AB was positive; all other combinations of A, B, and X were negative.

Pigeons were able to discriminate positive from negative sequences after training, pecking more to positive than to negative sequences. They discriminated more easily sequences ending in the incorrect stimulus than sequences ending in the correct stimulus. Most important for present purposes, however, is Experiment 3a in the series described by Weisman et al. (1980). In this experiment a sequence of two colors was followed by the presentation of a line orientation stimulus. One sequence of two colors was positive for one line orientation, with pecks to the line stimulus reinforced, and it was negative for the other. The other sequence of two colors was negative for the first line orientation and positive for the second. The two other sequences, repeats of each color, were negative for both line orientations. The two color stimuli were presented in sequence for approximately 2 seconds each, separated and followed by a .5 second ISI. The rate of pecking to the first 5 seconds of the line orientation stimuli was measured.

Acquisition of this conditional sequence procedure was about as rapid as the acquisition of the simpler sequential procedure. The pigeons readily discriminated positive sequences from others involving the same stimuli, indicating that the birds could maintain the information about

the joint occurrence of the two nominal stimuli and their order.

Appropriate performance of a sequence discrimination task requires the subject to maintain in reference memory general information determining the values of the different sequences; and to maintain particular information in working memory regarding the sequence actually presented on the current trial. On each trial the subject's task is to compare the information contained in working memory with the information contained in reference memory, and, on the basis of that comparison, to decide whether or not to respond. I now turn to a consideration of some of the schemes pigeons could use to represent the information in working memory.

For the most part, theories of representation for sequence processing in humans tend to be retrospective. That is, at the end of the sequence the subject is assumed to look back at the sequence that was presented, examine it for appropriate characteristics, and to perform accordingly (e.g. Sternberg 1975). This view probably follows from the usual techniques used to measure human performance, such as recognition or recall. According to associative models, during acquisition of a sequence, subjects associate items with their serial position either directly or through interitem associations, which may include "control elements" (Estes 1972). Another view of human sequence processing is based on the assumption that item information is stored separately and independently from order information (Murdock 1976). Because the experiments used to investigate these issues in humans typically involve the presentation of a novel list followed by a relatively immediate test for that list, they place a heavy load on the subject's working memory (cf. Sands & Wright 1980; Devine et al. 1979).

In the usual sequence discrimination experiment with animals, however, the stimuli used on each trial are typically well-learned and the working memory load can potentially be much lower. For example, in the first two experiments presented by Weisman et al. (1980), positive trials were determined by a two-stimulus sequence followed by a neutral test stimulus. The subject *could* perform the task as follows: Attend to the first stimulus; if it is correct, remember "peck." If while you are remembering "peck," the correct stimulus appears in position two, continue to remember "peck," otherwise remember "don't peck." When the neutral test stimulus appears, peck if remembering "peck." With this scheme there is no requirement that the information held in working memory, sufficient for deciding the value of the sequence, take the form of analogs of the stimuli presented and of their order. Instead, the subject could form intermediate decisions about the sequence and maintain those decisions. In the case of two-element sequences the decision is sufficient to specify the ordinal position and no further order information need be maintained. In many cases this decision memory would greatly reduce the number of items that need to be represented (one decision vs. two stimuli).

Alternatively, the information about stimuli and their order could be coded (a) as analogs of the particular stimuli and their order, (b) as analogs of the stimuli alone, but in a manner that allows them to vary along some dimension – such as strength with time since the presentation, (c) in the form of a running decision about the

appropriateness of the observed stimuli, and the next appropriate stimulus, or (d) as some combination of these three, changing from one form to another during the course of the trial.

The first two schemes are retrospective. At the time the test stimulus is presented the subject compares the representation in working memory with the representation in reference memory and makes a decision. The third model, however, is prospective (Honig & Wasserman 1981). At each point in the sequence the subject consults reference memory, generates a prediction regarding the next appropriate stimulus, and compares the presented stimulus with that prediction. This prospective model is compatible with all of the sequence discrimination data presented so far (but not with the production and probe data, e.g. Devine et al. 1979). Perhaps its application would be most instructive in the case of the conditional discrimination procedure presented by Weisman et al. (1980) because the most intuitive model in that case is retrospective, suggesting that the subject decides during presentation of the line orientation test stimulus whether the appropriate sequence had been presented. The prospective model, on the other hand, suggests that when the first stimulus is presented, the subject generates a prediction about the next stimulus that would appear in a reinforced sequence. If the first stimulus is "A" in that experiment, then only stimulus "B" would be appropriate for a positive sequence. Similar logic applies to the transition from the second element to the test stimulus. When "B" is presented as expected, a prediction of "horizontal" is generated, etc. This model requires that at each stage the currently presented stimulus be identified and that a prediction of the next appropriate stimulus be generated, conditional on the consistency of the present stimulus with a reinforced sequence. Once a negative decision has been reached, no further predictions need be generated.

The prospective model assumes that the memory load remains constant throughout a sequence whereas a retrospective model must assume that memory load increases until a decision is made. In a test of these models, pigeons were trained in a two-event sequence discrimination task (color sequence AB was positive) (Weisman & DiFranco 1981a). Choice accuracy varied as a function of the interstimulus interval (ISI) and retention interval (RI), and stimulus durations. With long ISI (8 seconds) and short initial stimulus durations (1 second), choice accuracy dropped with all sequences ending in the correct stimulus, B. On the other hand, increasing the duration of the RI from 1 to 8 seconds had little effect. Finally, presenting both stimuli for 1 second each caused choice accuracy to decline on all sequences, including those containing the dummy stimulus "X."

The differential effects of ISI and RI durations on performance and the different effects of short first-stimuli as opposed to short first- and short second-stimuli suggest that some kind of recoding is occurring during the course of the sequence. According to the prospective model, the bird uses the first stimulus to generate a prediction of the expected second stimulus, remembering that prediction (in working memory) until the second is presented, when it makes a decision about the value of the sequence. Longer initial stimuli are necessary because more processing must go on to generate a prediction than to

confirm one. Retention interval durations are less damaging than ISI durations because the presentation of the two nominal stimuli in the sequence constitutes a conditional discrimination; the correctness of the second stimulus depends on the first stimulus presented. Once the second stimulus occurs, however, the task becomes a simple delayed discrimination, which is more easily remembered than is a conditional discrimination over long retention intervals (Honig & Wasserman 1981). Apparently it is easier to remember a decision (or a prediction of food) than features of a stimulus.

Prospective schemes receive further support from studies of delayed matching-to-sample. Roitblat (1980) found that, as retention interval duration increased, pigeons were increasingly likely to confuse similar comparison stimuli as opposed to similar sample stimuli. Peterson and Trapold (1980) found that pigeons utilized the outcome of particular trial sequences as part of the representation. Correct choices on one sample-comparison pair were reinforced with access to food, the other with presentation of a tone. If, following a double reversal, the same sample-comparison pair was still correct but was followed by a disparate outcome, performance was initially good, quickly declined, and eventually recovered.

Aspects of a theory of representation in animal memory

In this section I return to the general problem of representation in light of the data reviewed above and with particular attention to the representations used in animal memory. I attempt to provide some foundations for a general theory of representation in animal memory.

General framework. Animals, like humans, are most profitably viewed as active information processors. The alternative view, that all behavior is more or less directly elicited in a reflex-like manner by simple stimulus events, has been well-refuted (Gallistel 1980, Oatley 1978). Functional aspects of stimulation must be interpreted, processed, and sometimes stored. Animals use a wide variety of representations, serving a large number of different functions at different levels of complexity. This range includes the fairly simple representations involved in homeostatic mechanisms and circadian oscillators as well as the more complex rule-oriented modelling representations that take advantage of the regular, structural properties of the environment to organize information in spatiotemporal maps (Honig 1981).

Inclusion of such a wide variety of mechanisms under the rubric of representation serves two purposes. First, it illustrates that there is nothing fundamentally mysterious about the existence of representations. In fact, they are ubiquitous. Second, it illustrates the wide variety of functions that representations can serve. By virtue of their ability to model environmental phenomena, they provide a level of internal, structural complexity not admitted by exclusively reflex-based systems. For example, it may seem reasonable to describe delayed matching-to-sample as a task in which the sample elicits a response that is inhibited until the comparison stimuli are presented. This is a straightforward associationist explanation. Beyond the connection between the sample and

the choice response, no other internal structure is assumed; a new reflex has been instilled. The analysis of confusion errors (Roitblat 1980), however, clearly indicates that more structure is involved. More features of the experience are represented than simply the occurrence of the sample stimulus, including features of the comparison stimuli and the trial outcome (Peterson & Trapold 1980).

Medium. The role played by structural properties of the environment in cognitive representations is most apparent in tasks which place relatively little emphasis on remembering which trial is occurring (e.g. which sample, in DMTS), but which place heavier emphasis on properties of the task that do *not* differ from trial to trial or from day to day – that is, on tasks that emphasize the role of reference memory. In tasks such as Hulse and Dorsky's (1979) or Straub's (1979), the variations between trials are irrelevant to performance of the task; the same responses are made in the same order on each trial. In these cases the representation controlling performance can be stored in reference or long-term memory as a single concept or map. In other tasks, such as delayed matching-to-sample, performance depends critically on the particular stimuli occurring in that trial. In these situations, special training is necessary for animals to utilize such general concepts. Without this training the tendency appears to be (for pigeons, at least) to represent the factors controlling performance as a set of rules or means/ends relationships (Carter & Werner 1978; Peterson & Trapold 1980; Roitblat 1980; Wright, Santiago, Urciuoli & Sands 1981).

For the most part, animals' performance of the tasks described here has involved internal cognitive representations, although animals are clearly capable of using other, more external cues to represent events (e.g. Blough 1959; see also Riley, Cook & Lamb 1981). When external or behavioral representations are used, they function as reminders entered on a note pad. They are typically more durable (at least in tasks such as DMTS) and, if changed, the resulting choice behavior reflects that change. Although, strictly speaking, they are not memories, they are capable of representing the occurrence of the sample. Traditionally, students of animal memory have taken care to prevent the use of overt external mediators, but these may play an interesting role, and they are amenable to study.

Content. In tasks involving delayed discriminations (such as DMTS and delayed sequence discrimination) and those involving spatial memory, each event (e.g. a sample presentation or a visit to an arm) potentially results in three kinds of information. First, there is the information reflecting that one particular event, and not its alternative (e.g. sample "A" was presented). Second, there is information about the features of the event (e.g. properties such as hue, brightness, coincidental noise, etc., which may vary randomly or systematically from occurrence to occurrence). Third, there is information about the meaning of the event (e.g. what the next appropriate behavior is, what stimulus is signalled, etc.). Any one of the three kinds of information may be sufficient for the representation of that event. The working memory representations used by experienced animals performing the kinds of tasks described above appear to utilize mainly the first kind of information – that event "A" occurred – when

there is a list of items to remember, each of whose meaning is important (e.g. in spatial memory, Olton 1978; and in serial probe recognition tasks, Sands & Wright 1980 and Shimp 1976b). The third kind of information appears to be used when accurate performance can be achieved by remembering only one item at a time (e.g. Roitblat 1980; Weisman et al. 1980, Weisman & DiFranco 1981a). Prospective representations, when used, typically seem to maintain information about the meaning of the upcoming stimulus, coded as (or along with) its attributes (e.g. the correct comparison stimulus, Roitblat 1980) and the outcome of the appropriate behavior (Peterson & Trapold 1980). They appear to be preferred when the task can be solved by either prospective or retrospective analysis (Riley, Cook & Lamb 1981 present a more detailed discussion of these alternatives and of the factors controlling their use). Little is known about the maintained attributes of retrospective representations, but at least in the case of serial probe recognition tasks involving relatively novel stimuli (Sands & Wright 1980), attributes of those stimuli must be involved.

Codes. Representations do not convey information about an event simply because of some change in the medium. The information is an emergent property of the relationship between those changes and the events they represent, and it depends on the availability of processes embodying the appropriate transformation rules (i.e. the code) (Oatley 1978). The particular change may not, however, be irrelevant. Codes, like the events they represent, also have properties, and vary along dimensions. Like the digital codes used in computers, the properties of a code's dimensional variation may be orthogonal to the information they represent (the voltage level of "1" in a computer might be between +1.8 and +5 volts, but variations within that range are not related to the information being represented.) Other times, this variation is related to the information being coded, as in analog tape recorders in which an increase in the voltage level of the record corresponds to an increase in the amplitude of the stimulus. In this case it is said that the representation intrinsically models that particular feature (Palmer 1978). Differences among represented items correspond to analogous differences in the representation, in the same way that a resistor-capacitor circuit shares corresponding properties (e.g. damping oscillations) with an automobile's shock absorber and suspension (Oatley 1978).

The structure typically attributed to human memory codes, semantic networks, or proposition-systems (Bobrow 1975) mainly involves codes with orthogonal dimensional properties that extrinsically model those properties. The relationship between features of the representation and corresponding features of the represented item are arbitrary.

Dynamic properties of imaginal representations, such as temporal aspects of image rotation (Cooper 1975; Cooper & Shepard 1973) and image scanning (Kosslyn 1981), strongly suggest that intrinsically coded representations are involved in these tasks (time spent scanning or rotating an image depends on the amount of image scanned or the degree of rotation). That is, the mental representations underlying image processing have pic-

ture-like properties, not easily accommodated by representational systems in which those properties are arbitrarily coded by, for example, pointers to dimensional values or properties (Anderson 1978).

Similarly, the dynamic properties of animals' representational systems also argue for an intrinsic encoding. Recency effects, when lists must be remembered (Olton 1978; Shimp 1976b), suggest that those representations intrinsically code order. The changes in confusability seen over time in DMTS (Roitblat 1980) argue for intrinsically coded properties of the comparison stimuli (cf. Conrad 1964). The parallel between these representations is obvious. The codes involved in animal memory may not be images, but they do share properties with the stimuli they represent.

Conclusions

During the last couple of decades many of the fundamental assumptions of behaviorism have been found to be untenable (e.g. Bolles 1975). Although behavioral changes as a function of experience can probably always be described in the vocabulary of reflexology, such descriptions entail the use of covert, unobserved and perhaps unobservable interoceptive stimuli and responses such as r_g-s_g mechanisms. In principle, therefore, the assertion that stimulus-response mechanisms are inherently more accessible on an empirical level and simpler (i.e., more parsimonious) than representational conceptualizations, is unfounded.

The work cited above and elsewhere makes clear that new approaches involving information-processing concepts have encouraged psychologists to attack problems that would, if they were considered at all, be too unwieldy or intractable within the old stimulus-response paradigm. These experiments make it clear that using a representation concept does not require the use of introspection and sloppy methodology. Exacting tests of theories do not require isomorphic mappings of conceptual elements onto observable elements, but they do require the specification of precise questions whose answers can serve to reduce the range of potential alternative explanations. Investigations of these questions often rely on relatively subtle manipulations and measurements, involving factors more complicated than the percentage of correct choices made. For example, as part of their evidence, many of the studies described above made use of patterns of errors (Olton & Samuelson 1976; Roitblat 1980). The same system that produces correct responses also produces errors, so the pattern of errors can be very informative concerning the representational system.

The metatheory of representation outlined here specifies the kinds of information that are necessary for a complete description of a representational system. A fully developed metatheory can save much fruitless effort by helping researchers to avoid the pursuit of unanswerable questions. At the same time, a metatheory can serve as a heuristic for generating experiments, suggesting questions that might not otherwise be addressed.

The concept of representation is not suggested as an alternative to associationism in the sense that it forces the rejection of associationism. Instead, associationism is seen as one particular class of representation, insufficient

in itself to capture all factors involved in preserving the effects of experience. Something changes in an organism as a result of an experience. If one chooses to describe those changes simply as variations in stimulus properties or reinforcement histories, then one chooses to ignore a variety of interesting phenomena. At the same time, "having a representation" explains nothing. Rather, the information represented, the way in which it is represented, and its effect on behavior, holds the kernel of an explanation.

Adoption of this metatheory of representation is one of the indicators of a paradigmatic revolution (Kuhn 1970) away from the old behaviorist paradigm. Though the shortcomings of behaviorism have long been apparent, no sufficiently perspicuous substitute has been available. A cognitive approach based on representational and other information-processing mechanisms provides such an alternative. The cognitive approach is not so much a competitor to behaviorism as a radical reorganization of thinking, resulting in a new "way of seeing" (Kuhn 1970) the problems of behavior. This new framework should allow rapid progress in the understanding of behavioral and cognitive systems, not only in animals, but, one hopes, in humans as well.

ACKNOWLEDGMENT

Preparation of this manuscript was supported by NSF grant BNS 7914212. Request reprints from H. L. Roitblat, Department of Psychology, 406 Schermerhorn Hall, Columbia University, New York, NY 10027. I thank all my kind colleagues who commented on earlier drafts of this manuscript.

Open Peer Commentary

Commentaries submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

Misrepresenting behaviorism

Marc N. Branch

Psychology Department, University of Florida, Gainesville, Fla. 32611

Roitblat, in his brief review of studies on animal memory, does a good job of pointing out the inadequacies of a behavioristic position. Unfortunately, the type of behaviorism to which he directs his arguments has been dead for about 25 to 30 years. He, like many others, seems to equate behaviorism with Watsonian or Hullian reflex-based approaches. Modern behaviorism (at least the Skinnerian variety), by contrast, is not a restrictive stimulus-response-based system at all. One must wonder why it is that so many psychologists today do not appear to understand the tenets of modern behaviorism. For example, Roitblat shows either confusion or ignorance when he states that there are two ways to view animals: Either as "active information processors" or as units "[whose] behavior is more or less directly elicited in a reflex-like manner by simple stimulus events." I submit that there is a third way – the way a modern

behaviorist views animals. Specifically, an animal is viewed as a biological entity whose behavior is a function of the contingencies of survival that have operated on its ancestors and the contingencies between stimuli and behavior that occur during the animal's own lifetime.

Note that verbs of the past tense were used in the description of the behavioristic position: That is, "causes" of behavior are taken to be historical. This seems to be the point most often missed by those who confuse modern behaviorism with its reflex-oriented predecessors. From a behavioristic point of view "causes" produce their effects across a temporal gap, and in many cases that gap can be large. Thus, to be comfortable with a behavioristic position one must be comfortable with the notion of "action at a temporal distance" in much the same way that explanations in physics often rely on action at a spatial distance (e.g., in how magnetism is viewed).

Modern mentalistic (or cognitive) approaches to behavior as well as Watsonian and Hullian behaviorisms reflect a different kind of preference, a preference for "contiguous causality" or "connectionism." That is, both views demand that "causes" occur immediately before effects. Both kinds of behaviorism may be viewed, therefore, as espousing essentially a "billiard-ball psychology." Of course, to adhere to such a position is to adopt a nineteenth-century conception of science. The connectionistic approach is one that, at least in the area of cognitive psychology, often leads to an animistic view of the behaver as originator of his actions. The objection that modern behaviorism makes to such views is that preoccupation with supposedly autonomous actions on the part of the behaver (e.g., storing information, retrieving it, perceiving it, filtering it, etc.) makes it much less likely that we shall discover manipulable variables of which such alleged activities are a function. If processing is simply accepted as a given, as something that the subject does on his own, then we shall not discover variables that change the way that processing proceeds. A behavioristic approach, on the other hand, demands that we look to a subject's history for variables that determine how he perceives something, or whether or not it is stored (i.e., whether it influences behavior later).

The goals of science, as usually summarized, include prediction, interpretation, and control of nature. A cognitive approach chiefly emphasizes the first two of these. Roitblat presents several examples of sophisticated interpretation, none of which increases our ability to control the phenomena in question. For example, when discussing the performance of rats in radial-arm mazes, he states that to perform well in such tasks rats "must have access to information about the arms already chosen, or about arms where food remains." Given such an interpretation, do we now know how to improve a rat's performance in such mazes? Does saying that the rat has such a representation tell us anything that we did not already know about its behavior? I have yet to be convinced that postulating an internal structure on the basis of observed behavior, and then referring to that structure as a cause of the behavior, really adds a great deal to our ability to predict and control behavior.

As another example, when discussing Maier's (1929) experiments Roitblat argues that "general knowledge about the room was the main factor producing the differences between the novel and familiar room." This is precisely the kind of explanation to which a modern behaviorist most strongly objects. Note that the cause is an inaccessible mental (or cognitive) state. Nowhere are we told of the important factors leading to the knowledge that subsequently leads to the behavior. In other words, preoccupation with the inside story leads to ignoring potential manipulable variables. Of course, there is nothing to prevent one from seeking out the variables that produce the knowledge that subsequently produces the behavior, but once that has been accomplished the knowledge becomes superfluous, except that it, in some sense, bridges the temporal gap

and thus preserves the notion of contiguous causality. If one is not wedded to connectionism, then the assumption of knowledge is unparsimonious. One might argue that something changes in the animal that corresponds to knowledge, but that something is most assuredly physiological. Roitblat, specifically and correctly, denies that cognitive theorizing should be taken seriously as physiology.

To sum up, the major behaviorist objection to cognitive or mental explanations is not that they are not objective, or that they are based on introspection (although, as noted by Roitblat, that was a concern of Watson's), but rather that they take on the status of explanation in such a way as to decrease the probability of investigation of manipulable variables of which behavior is a function.

Roitblat argues that the cognitive approach may represent a "paradigm shift" in the Kuhnian sense (Kuhn 1970). If it does, then it is a unique instance in that it represents a regression rather than an advance to something new. Throughout recorded history explanations of behavior have rested mainly on internal (often autonomous) causes. Cognitive approaches to behavior represent merely an extension of this old tradition. We must ask ourselves how mankind has improved its lot by assuming mental determinants of behavior, and having done so it is not so difficult to adopt a position that emphasizes the discovery of manipulable variables of which behavior is a function.

Memory and rules in animal serial learning

E. J. Capaldi

*Department of Psychological Sciences, Purdue University, Lafayette, Ind.
47907*

Roitblat, in attempting to refute objections to the notion of representation, might have profited from an examination of Kendler's (1952) view that the question of what is learned, whether S-R associations, cognitive maps, or the like, is a theoretical blind alley, a meaningless pseudo-problem. Kendler argues, among other things, that hypotheses about what is learned are always empty because they are invariably based, not on informed argument, but on intuition and bias. Although the ultimate validity of this argument may be questioned, it nevertheless appears to contain a worthwhile caution: Claims as to what is represented should always be subject to careful scrutiny. In this connection Roitblat shoulders a heavy burden because his objective is to examine many tasks in order to arrive at a general theory of representation. Error in any one case (and each presents problems) might well undermine the general effort, or at least its credibility. Unfortunately, Roitblat has fallen into a trap that has gone unrecognized and has lured others before him, that of mistaking an apparently simple but incomplete serial learning hypothesis for a complete one. His various errors about what is represented in serial tasks may stem from this faulty analysis.

Provide rats in a runway with a series of food pellets (hedonic events) that decrease according to a weakly monotonic rule which you and the rat have learned. Knowing the rule, answer a question: On which trial of the series does the 0-pellet event occur? Of course you cannot possibly answer the question knowing only that the events describe a weakly monotonic series. To answer it, additional information is required, information the rat must possess (over and above the rule), because in (for instance) the weakly monotonic series 14-5-5-1-0 it has run slowly on the 0-pellet trial and rapidly and nondifferentially on the four earlier ones (Capaldi, Nawrocki & Verry, in press; Hulse & Dorsky 1977). What is required in addition to (or besides) the rule is some other representation(s), of which there

are many, such as associating each event with positional cues (e.g., Young 1962).

Hulse and his associates (e.g., Hulse & Dorsky 1977) have examined nonmonotonic, weakly monotonic, and monotonic series. They have concentrated exclusively on predicting differences between groups, suggesting that the simpler the rule structure of a series the faster it will be learned. Granting for the sake of argument the correctness of this rule-learning hypothesis, note that in and of itself it cannot explain how specific events come to be anticipated, for example, the 0-pellet event that produces slow running in the 14-5-5-1-0 series. However, no other anticipatory mechanism is suggested by Hulse and his associates. Capaldi and his associates (Capaldi, Verry & Davidson 1980a) have argued for a representational system in which events are encoded as stimulus-stimulus items, each hedonic event in the series being signaled by the memory of one or more previous hedonic events. The memory model can explain why each event in the series is anticipated, something the rule model is incapable of, and it can explain between-group differences, a major preoccupation of the Hulse model.

Roitblat argued that the memory model is more complex than the rule model because, in his words, it "requires the animal to maintain more information about the stimuli than is required by the rule-based model, which requires only that the stimuli be ordered." But if the organism knows only that the stimuli are ordered (for example, weakly monotonically), then, as we have seen, it has too little information to predict when any particular stimulus will occur. If Hulse's rule model provided the organism with the additional information it needed to anticipate specific events, then its simplicity would vanish and it might turn out to be even more complex than the memory model; or the additional assumptions entertained might make the rule representation superfluous; but, in any event, current predictions of the rule model might be modified or drastically altered.

Roitblat states that available evidence favors the rule model over the memory model, although as he puts it, "the critical experiments have not been run," and that contrary claims by Capaldi et al. (1980a) are in his words, "simply groundless." This formulation is unfortunate all around: It distorts an important issue in animal serial learning while undercutting Roitblat's objective of arriving at a general theory of animal representation. The question is, under what conditions do animals use rules in serial tasks and under what conditions do they use memory? Hulse (1980) suggests that rats employ rules when there is memory overload, rules being a mechanism for organizing otherwise intractable experience. Memory is not overloaded, according to Hulse, in paired-associates tasks and in certain cases of serial learning – say, in a series consisting of four events (see Capaldi & Molina 1979), or perhaps even in longer series if they are not presented repeatedly in the same day. Capaldi and his associates have argued that rats easily remember many hedonic events (see e.g., Capaldi 1966; Capaldi & Verry 1981) and that in any event memory is always a major anticipatory mechanism, not only in explicitly serial tasks but in all forms of instrumental learning (Capaldi 1966; Capaldi et al. 1980a; Hagg bloom 1980; Maki, Moe & Bierley 1977). Nor have Capaldi and his associates argued that animals do not use rules: They have argued only that a convincing case for this proposition has not yet been made (Capaldi et al. 1980a). It is recognized that the rule and memory formulations are not necessarily incompatible: A demonstration that rats use memory is not necessarily a demonstration that they do not use rules. In any event, it has recently been reported that rats use memory in a serial task in which Hulse explicitly predicts memory overload (Capaldi et al., in press).

ACKNOWLEDGMENT

Preparation of this commentary was supported in part by NSF Grant BNS 80-01171 to E. J. Capaldi.

Antimisrepresentationalism

A. Charles Catania

Department of Psychology, University of Maryland Baltimore County,
Catonsville, Md. 21228

Think of a cube, all six surfaces of which are painted red. Divide the cube into twenty-seven equal cubes by making two horizontal cuts and two sets of vertical cuts each. How many of the resulting cubes will have three faces painted red, how many two, how many one, and how many none? It is possible to solve this without seeing the cubes. . . . But the solution is easier if one can actually see the twenty-seven small cubes and count those of each kind. This is easiest in the presence of actual cubes, of course, and even a sketchy drawing will provide useful support, but many people solve the problem visually without visual stimulation (Skinner 1953, p. 273).

Few would disagree that the author of the cited passage is a behaviorist. One characteristic feature of the behaviorist position is the emphasis on the behavior of seeing rather than on what is seen. There is nothing of so-called S-R theory in the account. It does not grow out of connections or associations or r_g - s_g mechanisms. Rather, it grows out of the various functions of behavior, which include, among others, the effects of the consequences of responding (as in reinforcement and punishment) and the role of discriminative stimuli (as when a stimulus occasions a response because the response has certain consequences in its presence).

Roitblat criticizes the behavioral approach, but his account shows so little familiarity with the contemporary behavioral scene that at first my temptation was to refrain from commenting at all. On what literature does Roitblat base the claim that it is a "behaviorist assumption that the hypothetical, covert changes occurring in an organism during learning correspond to the overt physical changes that are observed"? Thorndike, who explicitly did not count himself as a behaviorist, is cited as an early theorist. Of those who came later, the only actual behaviorists seem to consist exclusively of Watson, Hull, and Spence, with no reference more recent than 1943, even though Roitblat claims that, by "the inclusion of covert stimuli and responses," the S-R formulation tried to accommodate 1975 results from behavioral contrast experiments on reinforcer magnitude. Tolman was no doubt mistaken when he called himself a behaviorist, for Roitblat has placed him in the other camp. Operant behavior is presumably also outside the limits of behaviorism if one is to believe Roitblat, who contrasts the representational view with its refuted alternative, "that all behavior is more or less directly elicited in a reflex-like manner by simple stimulus events." If behavioral accounts are only those that describe behavioral change "in the language of reflexology," whatever happened to emitted behavior? Never mind that Roitblat demonstrates no awareness of contemporary behavioral treatments of private events. The argument is neither new nor rare that behavioral accounts need not rest on S-R theory. It is therefore distressing to see the misrepresentation persist.

Sometimes criticisms of behavioral interpretations are well-informed. For example, Paivio (1975, p. 281) quotes from the Skinner passage cited above, and then sympathetically discusses the causal role of private events in such accounts. He correctly points out an apparent inconsistency. Skinner maintains on the one hand that mental events play no causal role in generating behavior but discusses on the other hand how private responding can create new stimuli that in turn cause subsequent behavior. Part of the problem is in Skinner's use of the term "mental": He reserves it for nonphysical events with dimensions different from the physical dimensions of behavior. Skinner's objection that mental events cannot be causal can sometimes be read simply as a rejection of traditional psycho-physical dualism. If one covertly carries a digit in adding a long column of figures, the carrying is for Skinner behavior rather

than mental stuff. Moreover, the history of the carrying is important, in that the covert behavior can be traced to its origins in the earlier overt forms it took when the carrying was being taught; its consequences are also important, in that the maintenance of the carrying will both determine and depend on the correctness of subsequent answers. Skinner (1966; 1969) provides several examples in which the solving of a problem consists of creating a discriminative stimulus, sometimes private, that then occasions appropriate behavior. The sequence of events is sufficiently explicit that the difficulty for some readers is probably more in Skinner's language of causation than in the phenomena he deals with.

So much for setting the behavioral record straight. Now, what about representations? I am not troubled to learn that organisms can behave on the basis of the spatial organization of the environment, especially when controlling features such as locations defined in terms of extramaze stimuli are so explicitly identified. I am not troubled to learn that organisms can discriminate on the basis of relational properties among environmental events, especially when, in matching, the "content of the representation appears to be the stimulus to which the animal should respond during the test phase, coded in a form isomorphic with the dimension of variation of the comparison stimuli" (or, in other words, when, in the stimulus properties relevant to the matching task, the dimensions of the internal representations are pretty much the same as those of the external stimuli). Nor am I troubled to learn that organisms can respond consistently both in discriminating and in producing sequences, especially when the order information in the analogs of particular stimuli is pretty much the same as that in the stimuli themselves. Contrary to Roitblat's assumptions, a behavioral account need not hold that response learning takes precedence over place learning, or that all relational discrimination must be reducible to simpler physical properties of individual stimuli, or that all organized sequential behavior is built up through chaining (Catania 1979, pp. 85-91, 131-37, 146, 156-60, and 337-47).

But for Roitblat, any change that mediates the effects of past experience, whatever its form or substance, is a representation. We are carried from feeding rates in mantids through cognitive maps, matching relations, and assorted stimulus or response sequences in rats and pigeons and monkeys. The only exclusions seem to be those in which something really exists outside the organism, such as representation in books and other external media. What we are not told is how a representation can lead to behavior. Theophrastus long ago stated the problem of the postulated inner copy, in a commentary on the sensory theories of Empedocles: "with regard to hearing, it is strange of him to imagine that he has really explained how creatures hear, when he has ascribed the process to internal sounds and assumed that the ear produces a sound within, like a bell. By means of this internal sound we might hear sounds without, but how should we hear this internal sound itself? The old problem would still confront us" (Stratton 1917, p. 85).

Roitblat concedes that, at least for the moment, we are not likely to find specific neurological or physiological substrates, and notes that it is "convenient to avoid being too concrete about the physical instantiation of cognitive representations." But then what difference does it make where the representation is? Instead of putting it inside the organism, we might as well put it back where it came from: Inside the human observer who discovered the spatial or relational or sequential property of the environment to which the organism was responding. This is consistent with Roitblat's point that "we are usually limited to describing codes by analogy to properties of observable stimuli and responses." The debate is perhaps not so much about whether representations exist as it is about which organism they belong in.

One more point can best be made by analogy. If you take a

piano apart to find the music, you will be disappointed. You may have just listened to a performance of the Waldstein sonata, but when you are done you will have only pieces of wire and felt and wood and so on. You will also be disappointed if you try to find behavior inside the organism. The environment plays upon us as the pianist plays upon the piano. Just as it may take different virtuosi to bring out what is special about pianos and what is special about harpsichords, what an environment brings out may depend on which organism it plays upon. Or, to take the analogy further, just as some players may bring out some music more skillfully from some instruments than from others, different environments may bring out different sorts of behavior from different organisms. But then where is the music represented? The argument is not incompatible with a cognitive account: "The organism may fruitfully be considered as a *theory of its environment*" (Weimer 1973).

But wait, you say, what about the player piano? In that case you can find a representation of the music: A roll of paper with holes punched into it. No matter that one can substitute, in more contemporary versions, a magnetic tape or a computer chip (this feature of the analogy was allowed for, in choosing a piano rather than some instrument less compatible with representational systems, such as a violin or a trumpet). The holes in the paper are not music, and though one can be concerned with the way in which the holes are translated into the action of the keyboard, the causal account is incomplete unless one can say how the holes got punched in the first place.

Paivio closed the paper cited earlier with the following passage: "Skinner warned us against the diversionary effects of fascination with the inner life. I agree that the possibility is omnipresent. Mentalistic ideas are so seductive that one is in danger of being led by them down the garden path of introspection and mysticism forever. For that reason, perhaps only a tough minded behaviorist can afford to entertain the seductress" (Paivio 1975, p. 287). Certainly Roitblat is not qualified.

Behaviorism's new cognitive representations: Paradigm regained

Arthur C. Danto

Department of Philosophy, Columbia University, New York, N.Y. 10027

The wealth of experimental data and the elegance of experimental design recorded in Roitblat's paper notwithstanding, I question whether we in fact have a new paradigm in the offing or merely a modification of the old behaviorist paradigm. So long as the form of representation remains unaltered [where the form in question is that of the stimulus-response (S-R) collocations insisted upon by behavioral theory] the behaviorist structures will prevail both in the spirit of research and in the shape in which the data are presented.

Consider Hunter's (1913) regulative principle: "If a selective response has been initiated by a certain stimulus, and if the response can still be made successfully in the absence of that stimulus, then the subject must be using something that functions for the stimulus in initiating and finding the correct response" (p. 2). "The point is clear," Roitblat adds: "The animals were using something to stand for the information about the correct door." The form determines the question of what plays the role of the stimulus, and in answering "a representation," Hunter – and Roitblat – are in effect interchanging one stimulus for another, in this instance something within the animal that stands for the same thing the old stimulus itself stood for, as though external and internal stimuli were coreferential. But this shows that the stimulus already had the function of a representation, to the degree at least that we construe "standing for something" as constituting representationality. "Standing

for" accordingly was already a component in the behaviorist paradigm.

For that matter, "standing for" has played a role in psychological theory from its empiricist beginnings, when "ideas" were the currency of available theory and the association of ideas the basic principle of psychological organization. One idea *i* would have been associated with another idea *j* when *i* stood for *j* and caused the subject to conceive of *j* upon the occurrence of *i* "without any demonstration of the necessity of their coexistence or our so much as knowing what makes them coexist" (Berkeley 1709, paragraph 25). Much the same denotational and causal connections are postulated to hold between stimulus and response, which merely externalize the (somewhat) discredited "ideas" of associationism, just as the movement of the subject being the externalization of the movement of the subject's mind. Ideas have always been representational vehicles, but so, for that matter, have stimuli. Having nothing necessary to do with the object of response, they have virtually the status of names, being arbitrarily (or conventionally) associated with their denotations. To the extent that constellations of names could count as a limited language, behaviorists have been instructing their put-upon mice and pigeons in a curious discourse.

My point then is that associationism and behaviorism share a representational paradigm in that both require a relatively weak semantic notion (denotation), and concern themselves with collocating arbitrarily conjoined events. To be sure, they are divided over an ideology of internal and external representations, but inasmuch as Roitblat's representations (ideas?) are likely to be neurally instantiated, it is difficult to give much scientific weight to the division. The skin cannot be an observationally impermeable, metaphysical membrane: The tradition behaviorists set themselves against is that of mind against body, not body against the world. Indeed, if a representation is only a modification of nervous tissue, then we need not even postulate a "language of thought" as a medium for this class of representations. [See Fodor: "Methodological Solipsism" *BBS* 3 (1) 1980.]

In my view, there will be a true revolution in animal psychology – a "paradigm shift," to speak with the vulgar – just when (and if) we are required to attribute inference to animals, to map inferences onto nervous tissue with something like practical syllogisms, and to treat conduct under the framework of rational action rather than as mere response to stimuli. For then representations, as bearers of truth-values, must be assumed to have something like sentential structure: They must not merely stand for states of the world but must be satisfied by states of the world. Then the form of psychological explanation will have to be greatly restructured, as it need not be under Roitblat's interpretation. The logical form required of such representations of representations (since intentionality must enter the picture) is only well enough understood for us to recognize that a first step has not been taken toward a satisfactory metatheory. Collaterally, the postulation of the sorts of representations Roitblat refers to is not really the much-hoped for first step of progress toward "the understanding of behavioral and cognitive systems . . . in humans as well [as animals]" – not unless we can (which I doubt) replace the structures of rational explanation with the semantically meager representations dealt with here. Roitblat's paper is more polemical than its author supposes, but not against the opponent he imagines.

Memory: A matter of fitness

Juan D. Delius

Experimentelle Tierpsychologie, Psychologisches Institut, Ruhr-Universität, 4630 Bochum, Federal Republic of Germany

Only one look through a microscope at a silver-stained section of a rat's or pigeon's brain is necessary to convince someone that

there must be more than a few Ss, arrows and Rs to the explanation of behavior. There is nothing mentalistic about the several million neurons times the several thousand synapses that make up the nervous system of these animals. Neither are they just decorative frills for the anatomist to wonder at; evolution, that harsh judge, will have seen to it that they are involved in some very useful, fitness-promoting operations. In any case, physiologists have not taken long to find out that the function of these neurons is information-processing. By any measure a pigeon's or rat's brain still far outclasses even the fanciest of man-made computers. These computers again show us that without large amounts of well-organized memory nothing much can be done to make them behave adaptively. When they do, they are inevitably said to model, to represent the environment on which they operate. It is inconceivable that biological computers can function differently. One cannot but agree with Roitblat that a representation of environmental and self-generated events in animal memory must now be accepted as an incontrovertible fact.

The theoretical framework that Roitblat presents by defining a number of terms is no doubt useful in providing an orderly series of conceptual pegs on which one can hang a fair range of empirical facts – as Roitblat does. Because these pegs are kept quite general one can hardly quarrel with them. The processing components, however, would seem to require more emphasis and differentiation, with at least a subdivision into a storage and a retrieval phase. When discussing concrete evidence Roitblat certainly plays down the intricacies of recall that are really the main practical source of indeterminacy when one wants to identify the contents of memory. Performance on a particular recall task may only reveal the minimal amount of information that is actually represented in memory.

This leads to the question of whether we can ever expect to proceed from a metatheory to a general theory of representation in memory. Take the case of the rats that collected food tidbits in a radial maze. If they were suddenly chased by a terrier, would other memory contents surface, perhaps, than if they had carried on feeding? Would one obtain the same recall performance if the arms of the maze were baited with large food-chunks – or water, or even sexual mates – instead of tidbits? More generally, memory operations have not evolved according to a grand, highly rational plan; they have come about by a tortuous, chance-driven historical process whose sole criterion is the maximization of gene survival. Any realistic theory of animal memory will have to pay due regard to arguments about optimal function in evolutionary terms.

Is Roitblat's metatheory perhaps too restricted by categories of human speech? It is somewhat equivocal to say that images are not pictures in the head but that it is practical to describe them in picture-like terms. For example, the properties of such a visual image with respect to retrieval processing may be significantly affected by whether they are stored in the frequency-phase or in the spatial correlation domain. A map may be easily distinguished from a list in the realm of semantics but that need not necessarily be so in the context of associative networks.

Could the metatheory not cramp the style of researchers? For example, to say that it is convenient in cognitive science to avoid being concrete about the physical instantiations of representations begs the question "Convenient to whom?" Let us be fair. Some of us cannot be bothered with understanding the physiology of neuronal synapses or the mathematics of associative networks, but assertions that they are provisionally irrelevant may be shortsighted. Rather, it seems likely that progress in the understanding of animal memory, as in the case of most other biological phenomena, will come about through a scientific free-for-all, catch-as-catch-can. After all, why should the brain's handling of information storage and readout be so fundamentally different in this respect from the genome's handling of information storage and readout? The latter we now positively

know to be a process that does not conform to any neat theory, although it doubtless performs an exceedingly neat function.

ACKNOWLEDGMENT

Supported by the Deutsche Forschungsgemeinschaft through its Sonderforschungsbereich 114.

The informational character of representations

Fred Dretske

Department of Philosophy, University of Wisconsin, Madison, Wisc. 53706

There is very little of what I understand in the target article with which I disagree. I think Roitblat overestimates the significance of positivistic objections to the concept of representation (e.g., I do not think the observational indeterminacy of this notion, any more than the observational indeterminacy of the concept of electron spin, is "potentially devastating" to successful theory building) – although this may be because I underestimate the amount of life left in the behavioristic paradigm. What bothers me, instead, is that there are some things, particularly about the general character of representations, that I do not understand. So let me take this opportunity to say what I think Roitblat must intend so as to elicit a clarifying assent or dissent.

In the section on the meaning of representation in animal memory Roitblat gives what is basically an information-theoretic description of representation. *R* (some structure, event, or state of affairs), represents *S* (source, stimulus, early experience or whatever) insofar as there is some degree of correspondence or isomorphism between the properties of *R* and the properties of *S*. The isomorphism may not be one-to-one, but *R* functions as a representation of *S* only insofar as it preserves or carries information about *S*; and it carries information about *S* only insofar as there is a mapping of the features of *S* onto the features of *R*. This mapping consists, presumably, of a network of lawlike dependencies between the features of *S* and the features of *R*.

Given this understanding, what Roitblat goes on to say about the various dimensions of representations seems to me to be both important and true. Two different structures can be representationally equivalent (i.e., they can carry the same information), properties of the representation need not resemble the properties they represent, and a representational system (in contrast to the representation itself) includes processing components that determine the system's interpretation of the representation. These facts about representations are especially important to keep in mind when thinking about cognitive representations because there is an unfortunate tendency to think that certain internal representations (e.g., images) must resemble what they represent or must be harnessed to some particular interpretative process.

So far so good. Roitblat goes on to say things about representations, however, that blur this picture. He refers to representations as remnants of previous experience that allow the experience to affect later behavior. After an experience has ceased, some change must endure to mediate subsequent effects of that experience and that change, we are told, whatever its form or substance, is what is meant by a representation.

Roitblat can mean anything he pleases by the word "representation," of course, but if he is going to mean two things then he had better be sure that these two things are really only different verbal expressions of one thing. In this case they are not. The second characterization of representation is (what I shall call) a causal characterization. This is not the same as the first, information-theoretic, definition.

To see why this is so, suppose event *C* modifies a structure in some way *M*. After *C* ceases, *M* persists and subsequently brings

about effect *E*. Think of *C* as some experience, *M* as a neural change produced by *C*, and *E* as a piece of subsequent behavior controlled by *M*. Is this change, *M*, a representation of *C* according to Roitblat's account? If we consult the second (causal) characterization, the answer would seem to be "yes." *M* is some change that endures beyond the expiration of its cause. It mediates later effects of *C* in the sense that *C* has its effects on *E* through the mediation of *M*. *M* is an intermediate link in this causal chain. So *M* is a representation of *C* – a remnant of previous experience that allows the experience to affect later behavior.

But if we look at the first characterization of representation, then the answer to the question, "Does *M* represent *C*?" is not clear. It isn't clear because we haven't yet been told whether there is an appropriate mapping of the properties of *C* onto the properties of *M*. That is, we haven't been told whether *M* carries information about *C* or, if it does, what information it carries. To illustrate this point, suppose I break your vase: Something I do causes your vase to break. The broken vase causes you distress. The change (broken vase) brought about by my act endures beyond my act to bring about certain effects (your distress). Although the broken vase is a causal mediator, it carries no information about its causal antecedents – at least it carries no information about who (or what) broke it. So, in the information-theoretical sense, the broken vase does not represent the act which brought it about (its cause). Neural changes mediating stimulus and response may (sometimes, at least) be like that.

Roitblat says that the existence of representations is a rather innocuous idea and should rarely be an issue for theoretical dispute. "If an organism's behavior can be shown empirically to be affected by past experience, then some representation of that experience *must* exist." The only reason I can think of for his saying this is that he is thinking of representations in the causal sense, as remnants of experience that control behavior. But the issue should be, not whether there are such remnants, but whether they carry information about the experiences that produced them. Only then are they representations according to Roitblat's metatheory of representation.

I harp on this point because I think there is some confusion among cognitive scientists (whether their favored terminology be "representation" or some other, more explicitly informational idiom) about whether they are talking about causal relations or something different. The tendency is to suppose that if *R* is some internal state of an organism aroused by the causal action of *X* on its receptor systems, and if *R* endures so as to eventually produce some response appropriate to the (earlier) presence of *X*, then *R* must represent *X* (carry information about *X*). This does not follow. *R* may, in fact, represent *X*, but facts about its causal role in the production of an appropriate response do not show that it does. What would show that it does would be facts about the mapping of *X* onto *R*.

Representation: A concept that fills no gaps

Robert Epstein

Department of Psychology and Social Relations, Harvard University, Cambridge, Mass. 02138

Four categories of complex behavioral phenomena have traditionally given behaviorists trouble, at least in part because they began studying them only recently: Novel behavior (which leads some to speak of "creativity"); complex, distinctively human behavior not easily traceable to genes or the environment (which leads some to speak of such things as "the self-concept" or a "language organ"); covert behavior ("ideas," "thoughts," "percepts," "feelings," and so on); and behavior under the control of events that occurred in the remote past. Roitblat is

troubled by the latter category and appeals to such notions as "memory" and "representation" in his attempt to deal with it.

A pigeon pecks a plastic key that is transilluminated with yellow light. The light is extinguished, and 5 seconds later two other keys are illuminated – one yellow and the other green. The pigeon pecks the yellow one. How is it that the pigeon is able to do this? The answer, unfortunately – except in cases in which there is obvious mediating behavior (e.g., Blough 1959; Epstein & Skinner 1981) – is that we don't yet know.

The term "representation" sheds no light on the problem. It is one of a large number of cognitive terms in wide use that seem to fill a gap where the facts are not yet available. As Roitblat himself notes: We in no way explain the pigeon's behavior by saying that it "has a representation" of yellow, and alternatives such as "the pigeon was able to do this because it has a memory" or "because it remembers" are also uninformative. Having seen the sample stimulus, the pigeon was changed in some way, presumably physical. Neurophysiology is as yet too crude a science to identify the change, but it is surely there to be found.

Roitblat believes that we can make inferences about such changes from behavioral data, and that is surely true. Biologists and psychologists who study the biological basis of behavior do so routinely. But Roitblat's models are not physiological or anatomical, and in fact he presents no biological data whatsoever. Rather, Roitblat, like many cognitivists, uses behavioral data to construct models of information-processing systems that might generate similar data. He is not telling us what's inside, but rather how a computer might simulate behavior. The validity of this enterprise rests on the debatable assertion that organisms really are information-processing systems. Even Newell and Simon (1972) note that this is an assertion, but somehow an analysis of it gets lost among their myriads of models. I have examined this issue in some detail elsewhere (Epstein 1981) and here will note only that if the assertion is wrong, then so, most likely, are Roitblat's models of representation.

Roitblat gives undue weight to this cognitive construct with some parentheses and a slash. Speaking about the rate at which foraging animals obtain food, he writes, "In some animals this information is undoubtedly represented neurally (cognitively)." Does the punctuation signify that "neurally" and "cognitively" are one and the same thing? Later, he writes that the medium of a representation "appears to be neural/cognitive." Does this mean both "neural" and "cognitive," which are different things? In spite of these juxtapositions, it is clear from the paper as a whole that the two terms are not meant to be synonyms. Representation has no neural status for Roitblat, though he seems to wish that it did.

Many facts are cited about how behavior changes as a function of species, experience, and current circumstances. The facts speak for themselves and are the basis of a powerful heuristic. Roitblat's construct tends to divert attention from them and hence impedes a fuller understanding of their contributions. The sentence "experienced birds use some kind of map . . . to control their flight paths," in the context in which it is used, means only that, adult birds, after being displaced from their migratory routes, find their way back more successfully than young birds (What aspects of experience are important? Is age critical? Is maturation a factor? What properties of a bird's current environment control its return to the migratory route?). The statement "animals use experience-derived models of their environment to control behavior" tells us nothing more than that experienced animals negotiate their environments more successfully than unexperienced ones (What are the relevant experiences?).

Roitblat's concept is no substitute for facts about how genes and the environment determine the behavior he attributes to it, or for facts about how changes in behavior are mediated by the body. The concept may impede the search for such facts. He has

not shown, furthermore, that a model of representation will lead to discoveries about the nervous system. Though he admits that representation is not an explanation of behavior, he asserts more than once that models of representation will provide the "kernel of an explanation." I fail to see the difference.

Roitblat's whipping boy, associationist behaviorism, is not the only kind of behaviorism. Skinner's many detailed analyses (e.g., 1945, 1957, 1977) of the covert processes that Roitblat subsumes under the term "representation" are conveniently omitted. Why Roitblat even bothers to discuss behaviorism is not clear; and, indeed, the fact that his concept is presented in so doctrinaire a fashion makes it all the more suspect.

Some distinctions among representations

M. Gopnik

Department of Linguistics, McGill University, Montreal, Quebec, Canada
H3A 2K6

Roitblat makes a convincing argument that some instances of animal behavior are best explained by positing a representational system. He marshalls the evidence for the necessity of such systems in a careful and thorough way. However, it may be useful here to draw some further distinctions in the metatheory of representations.

Although it may be true that we need representation to explain how previous experience affects later behavior, it doesn't follow that the same kinds of representations are involved in all cases in which earlier experiences affect later behavior. In particular, it seems best to distinguish representations in which there is some direct causal connection between the states from those in which there is no direct causal connection. For example, it seems too broad an extension of the term to say that "Mantids represent their feeding rate by the volume of their gut contents [emphasis added]" In the normal state of mantids the volume of their gut content is a direct physiological result of their feeding rate. It is interesting, no doubt, that some physiological consequences function as information for later action and some physiological consequences do not. I get hungry when, among other things, my blood sugar reaches a certain level. [See LeMagnen: "The Metabolic Basis of Dual Periodicity of Feeding in Rats" *BBS* 5 (1) 1982.] There are, I suppose, other physiological changes that are a direct result of digestion rates and do not "tell me" that I'm hungry. But I am uncomfortable saying that the representation of my hunger by my blood sugar level is similar to the representation of my route home by some map of the neighborhood. It is like the difference between saying that every inch on a graph represents a degree of temperature and saying that every inch on a column of mercury represents a degree of temperature. In the first case the connection is arbitrary. The graph preserves the information about temperature but the form of the graph is not caused by the temperature. In the thermometer the height of mercury preserves the information about temperature because a change in its height is caused by a change in temperature.

Roitblat is right when he says that no one objects to representations if they are used in this causal sense. We may have to postulate a kind of internal Rube-Goldberg device that turns actions into enzymes into acids into actions again, but as long as the steps can be shown to be causally related nobody objects to actions causing actions through the mediating "representations" of enzymes and acids. Roitblat tries to argue that my knowing I'm hungry and my knowing I'm home are both similarly mediated by representations. The first is coded in blood chemistry and the second is coded in "a neural medium." But that is not really what cognitive representations are like, or what they are about. Roitblat admits that "it is so far more convenient to avoid being too concrete about the physical instantiation of cognitive representations." We are sure that something must be

going on with our neurons, and it probably has something to do with biochemical processes, but all the rest is handwaving. We are not really looking for a brain state to correspond to my thinking about (say) my grandmother.

In a larger sense, however, it really doesn't matter. We should think of cognitive representations as graphs rather than as thermometers. They are abstract models of some set of features of the world which we must postulate as serving in some mediating capacity in order to account for the complex activity of organisms. If we want to describe simple systems then we may be able to relate inputs to outputs directly without postulating any abstract mediating mechanism. But, as Roitblat convincingly shows, if we want to describe more complex systems, which include, for example, the ways in which rats solve problems and the patterns and kinds of errors they make, then the most economical description is one which postulates abstract representations.

Roitblat does draw a distinction between intrinsically and extrinsically coded representations and claims that the dynamic properties of animals' representational systems argue for their being intrinsically coded. But the distinction he draws between intrinsic and extrinsic coding has to do with whether differences among represented items correspond to analogous differences in the representations. A causal representation may be intrinsically or extrinsically coded; a noncausal representation may likewise be intrinsically or extrinsically coded. An example might be useful here. Suppose I want to construct a representation of a symphony. I might make a record of it or I might make a laser disc of it; I might hum its themes or I might write out the score. In the first two cases the sound waves of the symphony produce vibrations that cause electrical currents, which, in turn, cause a needle to cut grooves of a particular shape or particular digital information to be recorded (I here confess that I haven't any idea about the internal workings of records or laser discs, but I am confident that they are causal, not magical). My humming would preserve many of its dynamic properties, like duration and smooth transitions from one tone to another, but in no way is my humming caused by the symphony except that I intend that my humming should be like the symphony. This is what Roitblat thinks imaginal representations and animal representations must be like (without the necessity of conscious intention, of course). The score is neither caused by the symphony nor does it share with the symphony what Roitblat calls "dynamic properties." A half note does not take longer to read than an eighth note and there is no direct relation between the time it takes to read a portion of the score and the time it takes to play that part of the symphony. (We might note that no one of these systems is intrinsically better than another at representing information.)

What is clear is that noncausal as well as causal representations do allow us to account for complex behavior, and because different representations predict different kinds of behavior, we can test these consequences by looking at these differences. If we have a causal representation, then we have a more or less direct way of checking whether a particular experience results in a specific state and whether that state, in turn, affects future behavior. We can, for instance, induce the state artificially and see whether the predicted behavior results. However, in non-causal representations there is no such specified internal state. The representation is an abstract model that economically accounts for all of the data. We can fill a mantid's gut and see if his feeding rate changes, but we cannot artificially supply a rat with a representation of a maze and see if he behaves appropriately. We can only supply the rat with different kinds of experiences that we know require different representations and see whether he behaves in appropriately different ways. Causal and non-causal representations are therefore very different in the ways in which we can test them, and therefore a claim for their fundamental similarity and continuity seems unwarranted.

That, however, does not constitute an argument against using

noncausal representations. It just means that we cannot legitimately import our comfort with causal representations into explanations which use noncausal representations. Our explanations in terms of noncausal representations can show that the representation provides an economical and insightful account of the full range of behavior and that the animal (or machine, for that matter) behaves as if it were using such a representational system. New experiments can help us refine our description of the properties of the representational system, but they cannot demonstrate that the organism really uses that system. (There are some arguments that can be used to persuade us that a creature which seems to have an inner life does indeed have one, but they are philosophical arguments that do not produce the kinds of evidence or use the principles of reasoning which Roitblat is talking about. To go into them in detail here would take us too far afield.) But true noncausal representations are not in any worse shape than behaviorist models. The behaviorists simply constructed explanations which claimed that the organism behaved as if it had no internal processing (when they didn't sneak in covert mechanisms which really entailed internal processing). As Roitblat has shown, there are some kinds of behavior that simply cannot be accounted for under this assumption. New and more complex data can, however, be accounted for by postulating representational systems as abstract explanatory concepts. These representational systems, although themselves abstract models, have consequences in observable behavior. They can therefore be tested in the same ways any other abstract model that predicts behavior can be tested. The postulation of representation systems explains more data, leads to new, clearly-defined problems, and further integrates former partial explanations; for these reasons, it should be applauded and pursued.

Comparative cognition revisited

Stewart H. Hulse

Department of Psychology, Johns Hopkins University, Baltimore, Md.
21218

In the preface to a book published in 1978, *Cognitive Processes in Animal Behavior*, Harry Fowler, Werner Honig and I made a remark to the effect that our volume, among other things, perhaps symbolized the renaissance in experimental psychology of the study of the "higher mental functions" of animals. We suggested that "comparative cognition" might be an apt label for this fresh perspective, and we noted that the empirical techniques and theoretical strategies at work were often borrowed from developments in the study of human information-processing. Roitblat's discussion of the concept of representation seems to me to constitute a status report on the maturation of comparative cognition. All the hallmarks as we saw them in 1978 are there: The metatheoretical approach based on a model for representation, developed with human cognitive function in mind (Palmer 1978); the strong emphasis on memory as a keystone for representation; and, therefore, a choice of processes for empirical discussion and development that are all heavily loaded with descriptions of the maintenance of information across time. Roitblat's clear and interesting exposition indicates that the study of information-processing in nonhuman animals is unfolding quite nicely.

As we also noted in 1978, there has been a shift in the paradigms guiding research with humans and nonhumans. Whereas the models for human memory that guided research in the 1940s and early 1950s were based on developments in animal laboratories (and, fundamentally, on a general process theory of learning), the advent of the computer metaphor led students of human behavior to adopt a fresh theoretical path guided by the concept of information-processing, broadly defined. What we acknowledged explicitly in 1978 was that stu-

dents of animal learning and behavior, although keeping biological constraints in mind, were reshaping their research programs to capitalize on those information-processing models. In a very real sense, the paradigmatic shoe was now on the other foot, and comparative cognition was being stimulated by a search in animal behavior for processes and functional systems analogous (if not homologous) to those found in human behavior. Once again, Roitblat's article is a useful sampling of the important progress that has been made, especially since the late 1970s.

But Roitblat's discussion also testifies to the catch-up game comparative cognitivists have been playing. I doubt, for example, that students of human information-processing will find very much in the paper that will, to shuffle shoes again, facilitate research or theoretical development in their field. Nowhere, for example, do I see any data or theory that will urge students of human memory to adopt new ways of thinking. Nor do I see facts or theories that challenge conceptions of human information-processing so as to invite rebuttal and thus stimulate new research. Roitblat neglects, for example, all the work on the putative development of language in chimpanzees which, if such language exists, surely constitutes representation in animals in one of its more elegant forms. [See *BBS special Cognition and Consciousness in Nonhuman Species*, *BBS* 1 (4) 1978.] This is now of course a vast field, and a comparison of the work of the Gardners, Premack, Terrace, Sebeok, the Rumbaughs, and so on would not only lead to voluminous discussion but might also speak more to the controversy attached to the problem than anything else. But to my mind one of the most important things to come out of that controversy was not the issue of whether chimps have language, but the fact that some serious work was stimulated, refining and sharpening our ideas of the fundamental defining characteristics of language itself, human or otherwise. Here, potentially, is comparative cognition at its very best. Roitblat could, for example, have touched on that issue or, at least, on one of the important principles it emphasized, in his discussion of sequential representations. As Lashley (1951) pointed out many years ago, an understanding of the properties of syntactic serial structures is fundamental to many psychological processes, language (or music) in particular. Roitblat gives us a nice discussion of current knowledge about the processing of serially-organized information in animals, but he neglects related work based on analyses of language-like behavior (e.g., Thompson & Church 1980), and he certainly doesn't tell us why any of the work on sequential representation might be important in some larger sense.

My commentary is open to the argument that I have taken Roitblat to task on grounds that are irrelevant to the main thrust of his article. After all, his title restricts the discussion of representation to animals. Also, I have been critical because Roitblat has omitted topics that seem important and relevant to me, not because he has erred in handling the material he chose to develop. Perhaps there is injustice here somewhere. But it does seem to me that, by implication if nothing else, an appeal to concepts like representation couched in information-processing terms invites the direct comparative approach I have been stressing. To the extent that this is true, Roitblat's work, although an excellent development of the idea of representation by means of recent work with animals, did not go as far as it might have.

Representations as metaphors

Julian Jaynes

Department of Psychology, Princeton University, Princeton, N.J. 08544

We understand something by finding a metaphor for it with which we are more familiar, and the feeling of familiarity is the

feeling of understanding. In such metaphors, the thing to be understood is what I have called a metaphrand, and that more familiar thing to which it is compared, the metaphier (intentionally connoting multiplicand and multiplier in mathematics). This relationship between metaphrand and metaphier is at the basis of all knowledge: New phenomena are found to be similar to previous phenomena and are brought under their rubric (Jaynes 1976, Book I, chap. 2).

To say that in learning an animal stores the results of previous trials as a "representation," and that that representation controls future behavior, is to use representation as a metaphier. It seems innocuous enough. Moreover, it does indeed seem to help us understand what is going on by making the process familiar. But where does the familiarity come from?

As conscious human beings, we have embedded in us as a result of our culture and language the idea of mind as a kind of space, containing images, ideas, and memories that are representations of previous external events. This is what it is like to be introspectively conscious. But this is not what consciousness is. Moreover, so ingrained is our habit of imagining what is going on 'in' the minds of other human beings that we tend to carry over this "ejective perception" (as W. K. Clifford called it in the nineteenth century) to anything that behaves in any way similar to ourselves. So we wrongly impute consciousness to lower animals when what we are really doing is trying to know the animal predictively on the basis of past behavior. In other words, with terms like representation we may be using our own fallacious consciousness of consciousness as a metaphier to understand animals. This is a mistake.

Representationalism has, of course, become central to problems of artificial intelligence and is now influencing aspects of neuroscience. But I often suspect that hidden behind a great deal of sophistication are metaphors that are not always critically distinguished as such. Metaphiers such as 'information,' 'memory,' and even 'representation' have all but thrown off their psychological meaning. We may read that, in order to build a computer that can pass Turing's test of conversing in a conscious-like manner, we must build into it "a representational system . . . an active self-updating collection of structures organized to mirror the world as it evolves" – as if this abstract simplification were really what was going on in ourselves. It perhaps seems as if this were our consciousness, but careful examination of introspective experience shows that it is not (Jaynes 1976, Book I, chap. 1). Consciousness is primarily an analog 'I' 'narratizing' in a 'mind-space,' whose features are built up on the basis of metaphors. Present computer programs do not work on the basis of metaphors. Even if computers could simulate metaphoric processes, they still would not have the complex repertoire of physical behavioral activities over time to utilize as metaphiers to bring consciousness into being. Computers, therefore, are not – and cannot be – conscious.

Roitblat, in his full and interesting essay, feels it will turn the corner of an historical paradigm to take over the popularity of representationalism from AI into the field of animal learning. I certainly have sympathy for his purpose, that is, to once and for all put to rest the vacuousness of simple stimulus-response connectionism. But in this thorough interment, he is certainly stirring up old ghosts. The average reader of the term "representation" in the context of retention has as its first referent himself introspecting on an image, regardless of whether the theorist intends something nonpictorial like a matrix of propositions. Moreover, few would agree that it would be productive to use representation so widely – certainly not for the result or residue of any and all learning, as Roitblat suggests. Is something represented from the past in habituation? Or motor learning? Or in simple conditioning, as of leg flexion in the isolated ventral ganglion of a headless cockroach (Horridge 1963)?

If a dog, after searching a field, retrieves a stick that has

been thrown, we find it easy to say that he has some kind of tactal-olfactory-visual representation of the stick because he retrieves only the stick and rejects other objects. But perhaps we find this metaphier of representation from our own private experience easy in this instance because we know so little about such searching and retrieving.

Let us consider color vision, about which we know considerably more. We used to think that color was due to the wavelength of light reaching the eye, the ensuing sensation being a representation of the colored object. But simple color-shadow experiments (known, incidentally, since von Guericke's work in the 17th century) show that this is not true. If a red light and a white light are projected together onto the same screen, the result is, of course, pink. If we then put a hand in front of the white projector, we get a red shadow on a pink background. But if we instead occlude the shadow from the red projector, the outcome is not what would be predicted from the actual spectral composition. We do not get a whitish hand on a pink background (as we would expect on the basis of the ingrained representationalism of our consciousness of consciousness). Instead, the hand is a sharp blue-green hue (see Varela and Maturana 1981, from whom I have taken this example, for a much fuller discussion; also Land 1964). There is thus no one-to-one relationship between the wavelength of light from an object and the sensation. Color is experienced when there are certain ratios of relative activity in different color-specialized parts of the retinal field, and these ratios are weighted by comparisons with activity in the prestriate areas of the visual cortex and other neural surfaces. And when we take into account the further neuroanatomical evidence that, for each nerve fiber from a retinal ganglion cell entering the cortex via the lateral geniculate nucleus, there are a hundred or so fibers from other cortical and subcortical areas that enter at the same topographical location, we have a complexity in which the word "representation" seems lost indeed.

Feature detection work once seemed to lead to a complete representationalism. But these cellular responses have been shown to hold only under stabilized experimental conditions. Even simple perception, let alone animal memory, appears to be much more complicated and dynamic over time than the metaphiers of representation theory imply.

Memory representations in animals: Some metatheoretical issues

Roy Lachman and Janet L. Lachman

Department of Psychology and Bates College of Law, University of Houston, Houston, Tex. 77004

No single deficiency of psychological theorizing has invited more scathing criticism, and proved more devastating to the cumulative growth of knowledge, than the overgeneralization and overextension of theory beyond its appropriate context of usage (see Koch 1976). Generality and precision must be traded off in science; insofar as boundary conditions are not initially set or empirically determined, a given theory or "thema" (Holton 1975) is likely to be indiscriminately applied. The inevitable discouraging mismatches (between theoretically-derived prediction and data) at the level of precise empirical verification lead to a loss of credibility that may set into motion searches for a more productive reorientation.

Roitblat's effort to extend the notion of cognitive representation to theories of animal memory is the type of enterprise that, if done naively, can ultimately undermine the cognitive representational concept on which it relies. The global cognitive theories of the 1970s – for example, HAM (Anderson & Bower 1973), LAS (Anderson 1976), LNR (Norman & Rumelhart 1975),

and others – are extensions and elaborations of what are, at their core, theories of memory representation. These global models have been analyzed and described as “experiments in conceptualization” (Lachman, Lachman and Butterfield 1979), and as such, they increasingly appear to have failed. An immediate question, then, is this: If representational theory is not working well in the domain of human cognition, then how can it be expected to work as a theory of animal memory? Is Roitblat’s paper another instance of premature and inappropriate extension of psychological conceptualizations?

We think not. The analysis of cognitive representation Roitblat has chosen to extend to animal learning is that of Palmer (1978). Palmer’s paper does not present a theory. It is rather a metatheoretical analysis of the concept of representation that ranks among the most elegant analyses to be found in the cognitive literature. This analysis provides the most lucid interpretation of representation that can be applied logically to animal behavioral data, and Roitblat’s extension appears to be done with extreme care. It is calculated to meet numerous potential formal and empirical objections. For example, Roitblat recognizes that representational systems must be independently constructed for different species. He also understands the requirement of a task-by-task analysis, and acknowledges the potential problems involved in generalizing representational structures across task domains. Finally, Roitblat’s theorizing does not share the blinding abstractions of the global models or their highly unfavorable ratio of assumptions to predictions. Even if his effort is only partially successful, it should yield significant cumulative scientific progress.

A second issue concerns the kind of empirical and theoretical claim that is implicated when mental representational systems are attributed to infrahumans, and the kind of evidence that would constitute empirical corroboration of such claims. If memory representations in animals are considered a set of abstract mental structures that regulate observable behavior, then there are two classical views of their metatheoretical status: (1) They may be treated as theoretical idealizations, designed strictly to account for the corresponding empirical regularities (like, for example, the perfectly elastic molecules of the 19th-century kinetic theory of gases). The regularities accounted for are partly the result of a theorist’s choice of observations, and his judgments of their significance. Hence, a theorist may focus on invariance and develop a theory of representation that accommodates observable continuities across species and tasks while ignoring discontinuities. (2) Alternatively, such structures may entail an ontological claim. That is, the representations may be asserted to be entities with some kind of reality status similar to the molecules in the theory of Brownian motion. The second claim is the stronger one, because it posits some kind of real internal structures. These may be mental (i.e., the result of learning and experience) or they may be innate brain structures. In either case, the internal states and changes of state associated with such structures cannot be assumed to hold across species, populations, and task domains. Unlike the theorist asserting a weak form of representation, a theorist asserting a strong ontological claim will have no choice but to accommodate species and task discontinuities. Roitblat has acknowledged as much, suggesting by implication that he means to assert a strong ontological claim.

A frequent criticism of the strong claim for memory representation is that it resists direct empirical verification. Such verification has been demanded for the abstract internal states posited in the theories of Chomsky and of Piaget (see Feldman and Tolmin 1975), and can be expected in response to Roitblat’s theory of animal representation. The critics who call for direct empirical verification seldom, if ever, provide a clear account of what it is or what *a priori* it would constitute. This is not surprising; we have argued elsewhere that strong scientific claims are not ordinarily established by direct empirical verifica-

tion, but rather by a preponderance of evidence that the scientific community finds convincing (Lachman et al. 1979) Toward this end, we suggest the method of convergent validation described many years ago by Garner, Hake, and Eriksen (1956).

Cognitive psychology’s representation of behaviorism

A. W. Logue

Department of Psychology, State University of New York at Stony Brook, Stony Brook, N.Y. 11794

Many recent papers on cognitive psychology have taken behaviorism to task for attempting to explain all of behavior in terms of stimuli and responses. Roitblat’s target article makes a similar point. He demonstrates in case after case that a purely behavioral approach is inadequate in explaining many phenomena found in the animal psychology laboratory.

For example, Straub and Terrace (1982; see also Terrace & Straub 1981) have found that when pigeons trained to peck four colors in a certain sequence are presented with only some of the colors necessary to complete the sequence, they will peck the colors that are available in the correct order, simply leaving out the missing colors. If the pigeons had learned some sort of a simple behavioral chain in order to peck the colors in the correct sequence, then presenting the pigeons with fewer than the colors necessary to complete the sequence should have resulted in the production of sequences that terminated whenever a required color in the sequence was found missing. A more complex model is needed to explain these data and to predict new findings. Roitblat suggests that the pigeons learned to represent the sequence as an ordered set during training, and that they then scanned this representation while attempting to produce the sequence.

According to behaviorism, a truly scientific psychology can include only observable events, stimuli, and responses in its theories (Watson 1913) because only observable events can be objectively verified. This is what is taught about behaviorism; not usually taught in textbooks are the failures of this behaviorism. Roitblat points out two obvious ones: Hull’s (1943) resorting to interoceptive covert responses and stimuli in order to explain anticipatory processes, and Tolman’s (1948) postulation of cognitive maps in order to explain rats’ behavior in a maze. Both Hull and Tolman are classified by most historians as behaviorists, but these behaviorists were not able to make behaviorism work as it should.

Other failures of the behaviorists to hold to the behaviorist line are easy to find. For instance, when Watson (1970) had trouble explaining why the learning curves of humans were less regular than those of animals, he stated that “The human being becomes bored while learning. Other things stimulate him. The internal environment is complicated; internal speech (thought), for example, may always be a disturbing factor” (p. 211). Skinner (1953) found it useful to refer to a hypothetical hunger drive as a means of organizing data obtained from animals deprived to varying degrees. Even more recently behaviorists (e.g. Rachlin 1980) have constructed elaborate theories filled with many hypothetical constructs (for example “value”; Rachlin, Battalio, Kagel & Green 1981).

The fact is that, if you find a behaviorist, then you are also likely to find a cognitivist – or at least a part-time cognitivist (Logue, in press a, in press b). Roitblat and the cognitive psychologists have a point. Pure behaviorism is unsuccessful. No one has been able to take a strictly behavioristic viewpoint and explain all of the data. Although textbooks and students of the great behaviorists have interpreted behaviorism as accepting the existence and discussion of nothing but external stimuli

and responses, the great behaviorists were not themselves so imprudent.

It would be easy to conclude, then, that behaviorism has no message for today's cognitive psychology, and that in reality behaviorism and cognitive psychology are not all that different. However, although the cognitive approach does not actually represent as new and different a way of seeing the world as Roitblat and other cognitive psychologists have claimed (see, e.g., Miller, Galanter & Pribram 1960) there are several aspects of Roitblat's description of his metatheory of representation that appear to be at odds with the behaviorist approach.

First, when behaviorists have used hypothetical constructs, they have stressed that these constructs are hypothetical (e.g., Skinner 1953). However, in the target article, representations at times seem to take on a life of their own, to have independent properties that assert themselves in new situations rather than being abstract constructions. For example, in paragraph 3 of *Aspects of a theory of representation in animal memory*, Roitblat states "that there is nothing fundamentally mysterious about the existence of representations. In fact, they are ubiquitous. . . . By virtue of their ability to model environmental phenomena, they provide a level of internal, structural complexity not admitted by exclusively reflex-based systems." And although, in the last paragraph of "Representations and memory," he says that "no one is likely to argue that images are actually pictures in the head," in the first paragraph of "Animal navigation" he says "that experienced [migrating starlings] use some kind of map of their migratory route . . . to control their flight path." Actually, Roitblat's article does warn against thinking of models as real in several places, but because of the independent power that is still apparently ascribed to representations occasionally, it is easy to see how it could sound otherwise. Such errors can cause data to be misinterpreted and crucial experiments to be overlooked.

This brings us to the second point that those of the behaviorist school have emphasized about model building. Hypothetical constructs should be kept to an absolute minimum, not only because it is too easy to let them gain reality, but also because the closer scientists stay to their data the less likely it is that misinterpretations will occur (e.g., Watson 1924, chapter 1). Parsimonious models are to be preferred, and if there is no need for a hypothetical construct, then it should not be used. In the target article, for example, representations are discussed with respect to mantids. Mantids' foraging choices are apparently a function of their gut contents, an indirect measure of feeding rate. Roitblat therefore calls gut contents a noncognitive representation of feeding rate. But why bring in the concept of a representation at all? It has not been shown that anything is added by calling gut contents a representation; therefore, in this case, it is unnecessary. In the first paragraph of "Characteristics of representation," Roitblat states that "Representations can be seen to operate at many levels of behavioral functioning." A hypothetical construct may be very helpful in many places, but it should not be applied to widely varying cases without substantial evidence that this is essential, and most of the convincing examples in the present target article come only from learning and memory experiments in mammals.

Finally, a third danger often faced by a model containing many hypothetical constructs is that it can become untestable (Watson 1924, chapter 1). Roitblat addresses this issue when he discusses Anderson's (1978) objections to cognitive models employing representations. Anderson feels that, because one can always construct an alternative cognitive model that will make precisely the same predictions as a given model, it is really impossible to demonstrate that the given model is correct. Roitblat responded by claiming that only identical models would truly behave identically. However, it is clear from Roitblat's response that to test such an identity requires the ability to test many precise aspects of behavior, something that is usually beyond our current technology. Therefore, even if

Roitblat's theoretical argument is accepted, at a practical level Anderson's objections to cognitive models must stand. One also wonders why, if representations are purely hypothetical, it is necessary to worry so much about whether one or another representation is true, given each makes similar behavioral predictions.

Constructing cognitive models is a risky business. It appears that the behaviorists do have a valuable lesson to teach about this, but this does not mean that cognitive models must be abolished – indeed behaviorists themselves have depended greatly on cognitive mechanisms in the construction of their theories. Instead, the lesson is that restraint in using hypothetical constructs can help all psychologists towards better model construction.

A la représentation du temps perdu

John C. Marshall

Neuropsychology Unit, University Department of Clinical Neurology, The Radcliffe Infirmary, Oxford, OX2 6HE, England

When an animal's past experience affects its later behavior we can all agree that there must be something inside the animal's head that "mediates" the change. I understand radical behaviorism to involve the claim that such internal mediation can be fully described by specifying its neurophysiological instantiation: That is, the laws of stimulus-response (S-R) psychology purport to map directly onto an underlying neuronal mechanism. (That Skinnerians are not professionally interested in the physiology does not imply that they would expect decerebrated pigeons to peck response keys). In classical behaviorism a third estate is invoked, but, as this new intermediate level consists mainly in the replacement of capital letters by minuscule, rather little additional power is gained (Fodor 1965). I think then that Roitblat is entirely right when he argues that many interesting issues in animal memory simply cannot be addressed within these notational confines. His suggestion that learning and memory are better thought of in terms of an extended representational system involving the dynamics of change in domain, content, code, and medium is well-taken.

Roitblat urges us to adopt his metatheory of representations and then ask specific questions about the nature of particular memory systems in different species. I shall follow his advice.

With respect to human capacity, Bruner (1969) has drawn an important distinction between "memory with record and without record." Other students of cognition have preserved the dichotomy while changing the terminology. Thus Tulving (1972) speaks of "episodic" and "semantic" memory, O'Keefe and Nadel (1978) of "locale" and "taxon" memory, Marshall (1979) of "autobiographical" and "procedural" memory, and Eccles (1981) of "explicit" versus "implicit" memory. The crucial aspect of memory with record is that our representations include a time-labelled record of experience, organized so that we can coherently "summon up remembrance of things past." Bruner's distinction can be illustrated most sharply by pathology. In a well-known experiment, Claparède (1911) jabbed his amnestic patient's hand with a needle while shaking hands upon meeting. Moments later Claparède offered to shake hands again. The patient refused but seemed totally unaware of what had happened to her earlier. Many later studies of the (human) amnestic syndrome confirm that learning and memory – in the sense of experience affecting behavior – are possible without *rémembrance*. For example: Scoville and Milner's patient, H.M., learned a short visual maze after 155 trials and 265 errors; retested two years later, he showed 75% savings, reaching criterion in 39 trials and 69 errors, "although he did not remember the previous training sessions" (Milner 1970, p.

42). Similarly, an encephalitic patient reported by Starr and Phillips (1970) was taught to play a new tune on the piano; the following day, prompted by the first few bars, he gave a faultless rendition of the piece, but could not recall having learned it the previous afternoon.

Animals other than ourselves clearly display "memory." But do they have "memories" (as distinguished from mere feelings of familiarity)? And, if so, can they organize their memories into a structured autobiography? Gardner (1975) has argued that the assimilation of events into a history is facilitated by the semantic encoding available only to a language-using species. Even memories that reveal themselves as images must be hung onto a story-grammar if remembrance is to be distinguished from random hallucination (Freud 1895). If such speculations contain even a grain of truth, then it would seem that human memory requires a form of internal representation that is qualitatively distinct from animal memory.

Some thoughts on the proper foundations for the study of cognition in animals

Lynn Nadel

Cognitive Sciences Program, School of Social Sciences, University of California, Irvine, Calif. 92717

Roitblat has taken on a noble task. Precisely because of this, one wants to see the case for representation in animals stated as effectively as possible, but this I'm afraid he has not done. The major problem, which ramifies in several ways in his paper, is that he approaches cognitive functions as though they were as prepared for the demands of some laboratory task (any conceivable one) as they were for the basic regularities in the external world. Consider that as much of Roitblat's target article is devoted to the analysis of a rat's representation of sequences of rewards (which the rat does not do all that well), as is devoted to the rat's understanding and use of space (which it does extremely well). Does Roitblat believe that these two kinds of things are equivalent? That the rat in its natural habitat is as likely to call upon both, and hence devotes equal representational power to each?

It has been obvious for some time, as Roitblat points out, that some form of representational theory is needed to account for the facts of animal behavior. There can no longer be any doubt that internal models of various features in the environment underlie much of what we, as psychologists, seek to understand. Thus, to lay the foundations for the study of animal cognition is a crucial first step in the reconstruction of our theories of animal behavior. Such a step demands, in my view, at least two related sets of considerations. First, one wants to provide a metatheory of representation – and here Roitblat is generally on the right track. Second, one wants to provide a principled set of reasons for investigating some but not all possible representations. Here is where Roitblat's argument comes apart, and in so doing threatens his entire enterprise.

As long as one allows the nature of the tasks to define the kinds of representations one studies, there is little hope that the study of animal behavior from a cognitive perspective will accomplish much more than a relabeling of already-described phenomena. With some significant exceptions, much of what passes for cognitive psychology is the same old stimulus-response (S-R) stuff dressed up in new language – very often that of information-processing. Roitblat seems to understand that taking a cognitive stance amounts to more than this, yet he stumbles into the trap of letting operations and procedures pursued in the laboratory for the past several decades form the guidelines for his approach to cognition. This is a serious mistake.

To me, cognition refers to the possession of knowledge. The study of cognition focuses on the content of this knowledge, the nature of the system in which it is represented, the way in which it changes over time, the (physiological) means by which this representation and its changes through time are realized, and the ways in which the knowledge guides behavior. Not every kind of knowledge is equally likely to be the main theme of its own representational system. Thus, although it is quite reasonable to talk about representational systems for obvious environmental regularities (e.g., space and time), it makes little sense to assign equal status to representational systems modelling regularities in one or another artificial laboratory situation. Can anyone think of a convincing real-world regularity comparable to the reward-sequence learning Roitblat discusses, or indeed, the venerable delayed matching-to-sample task?

Central to Tolman's cognitive theory, naive as it might have been about performance issues, was the notion of "purpose." Animals form representations to serve purposes; representational systems evolve because they are useful in a species' ecological niche. What is missing in Roitblat's case is *biology*. Certain kinds of representational systems make biological sense, but other putative candidates seem to speak to no particular biological need. These should not be treated in the same way in our theories.

Let me illustrate what I mean by briefly considering Roitblat's treatment of spatial behavior in rats. Basing his remarks on recent work with the radial maze, he lets the characteristics of the task determine the nature of the representational systems. Instead of talking about spatial representations he focuses on the procedural distinction between "working" and "reference" memory. He assumes that these operational requirements actually tap into separate representational systems. Recent work, however, seems to indicate that working and reference memory in the radial maze are simply different states of activation within the same spatial representation system (Nadel & MacDonald 1980; Roberts & Dale 1981). Roitblat says that "the working memory representation consists of a list of spatial locations. As each choice is made, another spatial location is added to the list." But, what does such a list look like? What sort of medium would have the capacity to store transiently any kind of information? – perhaps a blank slate read by some unstated homunculus, or a tape recorder, but not a brain as I understand it.

Simple associative models of learning and memory, wisely discarded by Roitblat, have left behind the unprofitable presupposition that information from the external world impresses itself upon a system with no meaningful internal structure. One major thrust of cognitive analyses is that our hard-wired representational systems themselves contain information – knowledge embedded in their workings that constrains in important ways the kinds of information any given system can represent. The information inherent in a "spatial location" is not free to wander around the cognitive (neural) system, moving from one representational base to another. The information is part of a spatial representation, and cannot be meaningfully separated from this framework.

One might say that representational systems are basically interested in things (objects, events, etc.) as well as in the relations among things (spatial, temporal, causal). Among the tasks of the cognitive psychologist are the determination of what defines "thinghood," what sorts of relations are permissible, and how new things are created within the representational system. The tasks of students of animal cognition are much the same, and one hopes that the comparative/biological perspective of such students will help them focus on the kinds of representations truly indicative of a given species – and not on those addressed by arbitrary laboratory studies that are merely easy to do, or seductively successful, yet ultimately of little value in deciphering the complex biological structure of cognition.

But let me close on a more positive note. Roitblat has performed a major service by saying in public what most psychol-

ogists interested in animal behavior have been saying in private for some years. Behaviorism is dead. Long live cognitivism. Let us simply be sure that there is more in it than a name change.

On the content of representations

R. J. Nelson

Department of Philosophy, Case Western Reserve University, Cleveland, Ohio 44106

Roitblat attributes certain properties to animal representations that, it seems to me, cannot possibly obtain in the general case.

What I question is Roitblat's assertion that, for sufficiently complex animals, the relation of world to representational field is a "mapping," that content is constituted by anything at all like a "preservation of features," and that the mapping (or, more generally, "relation," if my criticism is right) is a "code" in the sense of a homomorphism from wholes of parts to wholes of parts.

Among representations in Roitblat's sense of remnants of experience that affect later behavior are representations of proximal stimuli. There is good reason to believe there is a two-step process involved here (Pylyshyn 1978, p. 20ff). The first is a cognition free, autonomous process during which input is actively transduced to form the representation. Let us call this step or stage "reception." The second step is "perception" in which the representation is identified and stored in working memory – tagged as a "type." This step is cognitive (Kolers & Perkins 1975).

In higher organisms (but probably not in frogs – see Lettvin, Maturana, McCulloch & Pitts 1959 – or even in somewhat higher forms) a wide variety of extremely fuzzy physical patterns might all be processed in reception as a single type of representation (the "universal" of traditional philosophy), a phenomenon largely attributable to the organism's expectations (Nelson 1975; 1976). Humans, at any rate, often take degraded input at one time to be a type T (i.e., they process input to form a representation of type T) and at another time to be T' (processing input to form T'), where T and T' are incompatible types or features. If this is indeed the case, then informational content cannot consist of the preservation of features (except in lower organisms such as frogs) except in some very partial, poorly understood manner, but must instead be contained in the computational structure of reception, that is, in attributes of process. (On abstract mathematical grounds it is well-known that radically different computational structures might process exactly the same input but lead to inequivalent computational states; see Davis 1958. If this is true of the *situs* of content, then *a fortiori* the "mapping" from a feature of the world to representation cannot be a coding.

It cannot even be a mapping, in the strict sense of a many-one relation. Consider, for example, gestalt phenomena: The situation here does not concern reception (as vagueness does), but rather perception, that is, recognition of a representation as of a type. Whatever else it is, a gestalt phenomenon is marked precisely by the interpretation of a single item of the domain as of two types. The same representation is cognitively processed in two ways, depending on contingencies that are at present inadequately understood. Thus whatever goes into working memory or beyond could be of one or two material types issuing from one object of the domain. Hence the formation path from the world to representations cannot be a mapping.

I do not think this analysis should be construed as an argument against the theory of representations, but rather as an argument in support of the view that representation depends essentially on active processing. Preserved content must be more fuzzy than is intimated by such concepts as coding or even mapping. The relations are many-many.

Premature closure of controversial issues concerning animal memory representations

William A. Roberts

Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5C2

Although I am certainly sympathetic toward the study of representations in animal memory, I must challenge some of Roitblat's conclusions. In several instances, I feel that he implies too much the closure of issues that should still be considered open.

My first concern is Roitblat's discussion of spatial representation in the rat in the radial maze. He suggests that there are two alternatives, a cognitive map or a list of places visited, and he argues that existing data favor the latter alternative. It strikes me that these two theories are not strongly orthogonal and have a number of attributes in common. In order to mark regions on its map as visited, a rat would probably need to partition the map into separate places. This procedure begins to come dangerously close to the notion of remembering a list of places. The functional consequences of a theory which holds that rats remember a list of places versus one which holds that they remember a segmented map may differ little and hence the distinction may be of little consequence.

Roitblat argues that the failure to find rats making spatial generalization errors in the radial maze argues against the map theory but is easily explained by the list theory. I don't think this is the case. In his criticism of the Suzuki, Augerinos and Black (1980) experiments, Roitblat points out quite correctly that rats running to the end of an alley on an elevated radial maze would not identify that arm just by the narrow set of cues at the end of the arm but by a wider set of cues taken in by a panoramic view. In this case, there should be overlap between the cues seen at the end of adjacent alleys and hence between their representations in the list.

Suppose that a rat had entered arms 1, 2, 3, 4, 6, 7, and 8 in the radial maze. When peering down arm 5, the rat would see a set of cues, some of them associated with the entered arms 4 and 6; when peering down arms 4 and 6, the rat would see sets of cues, some of them associated with the unentered arm 5. Given this assumption, it is easy to see that a list theory would predict spatial generalization, as would a map theory. Because rats do not show spatial generalization of errors around an unentered alley but return to the alleys entered earliest in a sequence, both theories may need additional assumptions. Roitblat's solution is to assume that the list of places visited preserves order information. One could just as well assume that the markers placed on the cognitive map include order or time tags. This assumption does not tell us, however, why a rat, remembering that a particular alley was the first one entered, would then choose this one to enter erroneously.

Actually, an assumption that rats' memories of places visited include order information is unnecessary. If it is assumed that rats forget memories of early alley entrances before memories of later alley entrances, as a result either of trace decay or of retroactive inhibition, then errors should be made mostly to alleys entered early because they will have the same status in memory as an unentered alley. This explanation works equally well with a list or map conception of representation.

A second issue I would like to debate is Roitblat's interpretation of his experiment on coding in pigeon short-term memory (Roitblat 1980). In his third experiment, pigeons were trained to perform delayed matching with three sample stimuli (colors) mapped onto three different comparison stimuli (line orientations). For example, a pigeon was reinforced for choosing a 90° line after a red sample, a 12.5° line after an orange sample, and a 0° line after a blue sample. On tests with the orange sample, more retention errors were made to the 0° line than to the 90° line, and a score derived from these data (the AB/BC distance

ratio) decreased as the delay between sample and test increased. Similar findings have been reported for monkeys by Gaffan (1977).

Roitblat's interpretation of this finding requires us to assume a double transformation of information on the part of the pigeon. Initially, the pigeon is assumed to code the orange sample as a 12.5° line at the time the orange stimulus was presented or shortly thereafter. Subsequently, during the retention interval, the 12.5° line becomes confused with the 0° line, and the 12.5° representation is transformed to a 0° representation. The tendency to confuse comparison stimuli increases with the retention interval because an increase in the retention interval provides extra time for the transformation from 12.5° to 0° to take place. This work is perceived as being analogous to Conrad's (1964) findings of acoustic confusions of visually-presented information in humans. However, the analogy does not seem complete. In Conrad's experiments, people were presented with a string of letters and in recall, they made substitution errors between letters with similar sounds. The confused items were initially placed in short-term memory. In Roitblat's experiment, the pigeon sees only the orange sample and then presumably codes it into a 12.5° representation. But where does the 0° representation it is to be confused with come from if its color code (blue) was not presented initially? We seem to be required to make still another assumption. That the pigeon retrieves a representation of the 0° line from long-term or reference memory in order to confuse it with the 12.5° representation.

These results may be explained with fewer assumptions about cognitive activities on the part of the pigeon. Let us assume that the pigeon encodes, if not a copy, then a representation of the sample stimulus that retains its salient attributes. Through previous training, the pigeon has learned to apply rules of choice at the time of test, such as "if orange, peck 12.5° line," "if blue, peck 0° line," and so on. If a pigeon remembers orange and, upon seeing the comparison stimuli, codes this memory into the instruction to peck the 12.5° line, then, by the principle of stimulus generalization, more errors should be made to the 0° line than to the 90° line. As the delay interval becomes longer, it is well-known that error frequency will increase, even if similar comparison stimuli are not used. If we assume that information about the orange sample stimulus, which remains in a progressively decaying memory, continues to bias errors toward the 0° line, then Roitblat's findings are predictable. That is, if the proportions of errors made to the 0° and 90° lines remain constant as total errors increase, the AB/BC distance ratio plotted by Roitblat should decline as the delay is increased. Therefore, although Roitblat's interpretation cannot be ruled out, a model assuming fewer cognitive transformations can handle these data.

Finally, I would argue that Roitblat presents an uneven account of the controversy between Hulse and Capaldi concerning the processing of serially-presented reward magnitudes. Roitblat concludes that the existing data favor Hulse's position and he dismisses Capaldi's arguments in a far too off-hand manner.

I am not sure what Roitblat means when he says that Capaldi's experiments "involve a different task from Hulse's, making it difficult to interpret them. Until it can be shown that their procedure can produce findings analogous to those obtained with Hulse's procedure, any assertions are simply groundless." Both Hulse and Capaldi use the same task, rewarded running in an alleyway. Hulse (1980) has argued that Capaldi's experiments differed from his in terms both of the total number and variety of reward transitions experienced by subjects and in the time intervals used between and within patterns of reward. Capaldi, Verry, and Davidson (1980b) have countered that these objections are either incorrect or inconsequential in the comparison of the two experiments. In any case, the important point made by Capaldi (Capaldi, Blitzer & Molina 1979; Capaldi & Molina 1979; Capaldi, Verry, & Davidson 1980a) is that there would be

no point to his doing experiments that either replicate Hulse's results or produce results analogous to them. In order to compare the rule-learning theory with the memory association theory, special designs must be used, and these lead to different predictions based on these theories.

Roitblat argues that Capaldi's theory predicts a better tracking of the 0 element in the weakly monotonic sequence, 14-5-5-1-0, than in the strongly monotonic series, 14-7-3-1-0, because the 1-pellet element gains less generalized signal strength in the former sequence. This conclusion highlights the confounding of sequential theory variables in Hulse's experiments that is emphasized by Capaldi. The fact that 1 is closer to 3 than to 5 would lead one to predict better tracking in the weakly monotonic series. However, the memory of 5 is associated with rewards of 5 and 1, whereas the memory of 3 is associated only with the reward of 1. Hence, the generalized signal value from 5 to 1 may be greater than that from 3 to 1, because 5 has a stronger signal value than 3.

In his simulation of predictions from Capaldi's theory, Roitblat finds that equivalent running speed should occur to the 3- and 0-pellet elements in an 18-3-6-10-0 sequence. But Hulse and Dorsky (1979) found much better tracking of the 0 element than of the 3-pellet element. Although this constitutes one bit of evidence against Capaldi's position, it does not seem overwhelming. It should be noted that Roitblat's simulation shows that the association theory predicts the major phenomenon of these experiments: Better tracking of the 0 element in the monotonic sequence than in the nonmonotonic sequence. Further, Roitblat completely fails to mention the experiments of Capaldi and his associates that argue against the rule-learning position. Instances are found in these experiments in which (1) tracking decreases when the number of rule instances is increased and (2) tracking is better with nonmonotonic sequences than with monotonic sequences. The outcomes of these experiments are in keeping with predictions from Capaldi's sequential theory.

The jury is not yet in on the question of what rats learn when they encounter a structured sequence of rewards. Hulse and his coworkers may yet provide a demonstration of rule-learning that cannot be dealt with by association principles. Or, Capaldi and his associates may convince us that association and generalization are the only principles necessary. At the present time, the issue is not as clearly decided as Roitblat contends.

The logic of representation

William W. Rozeeboom

Department of Psychology, University of Alberta, Edmonton, Alberta, Canada T6G 2E9

Do rats and pigeons act under the influence of representations? I don't know, and Roitblat has not convinced me that he does, either.

When organism o 's encounter with stimulus S at the time t_i makes a difference for what o later does at time t_k , evidently o 's total state at each intervening time t_j must include some condition R_j whose instantiation by o at t_j has been brought about by o 's experiencing S at t_i and which, in turn, affects o 's behavior at t_k . But that scarcely begins to imply that R_j "represents" S (or anything else) for o at t_j . Otherwise, any stage of any causal process, even one as simple as the chained falling of dominoes, would count as a representation of each stage that preceded it. To justify talk of behavioral "representations," it must be argued not only that the link between input and output is mediated by an internal condition of some complexity, but also that the mediator has the right kind of complexity both in function and in composition. Unfortunately, although intuitions about representation pervade our everyday thinking, such notions are still

far too obscure to provide insights into behavioral mechanisms. Theories about these mechanisms are much to be desired – but with representation figuring therein as explicandum, not as explicans.

We can surely agree that entity R is not a representation, unless there is some other thing T that R represents, or would represent were T to exist, under some mapping principle ρ that picks $T = \rho R$ out from all other aspirant representees. (Whether R 's ρ -relatum must always be strictly unique will not be discussed here.) But a representational system is not just any old triple $\langle R, \rho, T \rangle$ in which ρ is a function from domain $R = \{R_i\}$, onto range $T = \{T_j\}$, or else we should have to admit, for example, that telephone subscribers represent the numbers assigned to them by phone directories. Linguistic intuition demands more than that, something else that we intimate by saying that representations "depict," "symbolize," "signify," or "stand for" their objects (see any dictionary). For $\langle R, \rho, T \rangle$ to yield "depictions," complexes of R -units must correspond in some nontrivial fashion to certain complexes of T -things under element-to-element mapping ρ . And R_i symbolizes/signifies/stands for T_i only if it does so for some organism o at a time t by somehow standing proxy for T_i in o 's occurrent or dispositional functioning at t . But complexes of what sorts? And proxy how?

Three paradigms of representation can be discerned within established linguistic usage. Most basic is the primary cognitive "aboutness" (intentionality à la Brentano) that characterizes the contents of mental acts such as believing, perceiving, fearing, desiring, remembering, trying, and the like. Here, element domain R comprises concepts that purportedly designate (refer to, denote, are of) concrete or abstract objects, and form compounds whose psychonomic nature is still unknown,¹ most notably "propositions" which, when truthful, depict facts.

Second, when some suitably structured external or internal stimulus S_i seems capable of evoking in organism o an intentional content C_i that signifies some real or possible entity T_i distinct from S_i , we often say that S_i represents T_i for o , especially when, as in perception, the primary elicitor of C_i is T_i itself. Language is the example par excellence; but pictures, schematics, and most other stimulus patterns that common sense takes to encode "information" also illustrate well this concept of secondary aboutness. When S_i so represents T_i for o , it stands proxy for T_i in being a means for o to think about T_i in the latter's absence. Indeed, o 's percept of pattern S_i (or, if S_i is internal, S_i 's embodiment in o) may well correspond structurally to the composition of o 's concept C_i of T_i under an elicitational transduction of parts in such fashion that S_i can evoke C_i in o (i.e., depict T_i) even when o has never encountered T_i itself.² Such an S_i encodes information for o precisely to the extent that it makes available to o the propositional content of a belief.

Finally, modern mathematics formally views one relational system $\langle R, \phi \rangle$ as a representation of another, $\langle T, \psi \rangle$, under mapping ρ whenever (a) ϕ and ψ are binary (more generally n -ary) relations on respective domains R and T , (b) ρ is a function from R onto T , and (c) for all R_i, R_j in R , $\phi(R_i, R_j)$ holds if and only if $\psi(\rho R_i, \rho R_j)$ does. Despite superficial similarities, this is very different logically from cognitive representation, even though we have good reason to suspect that morphisms of ideation under causal mapping are important in the psychonomics of secondary aboutness. When $\langle R, \phi \rangle$ is isomorphic or homomorphic to $\langle T, \psi \rangle$ under ρ , and $\langle T, T_j \rangle = \langle \rho R_i, \rho R_j \rangle$, the fact $\phi(R_i, R_j)$ formally represents the fact $\psi(T_i, T_j)$ relative to mapping ρ but not relative to any particular organism o that this is a representation for; and $\phi(R_i, R_j)$ also simultaneously represents many other facts $\{\psi'(T'_i, T'_j)\}$ relative to other mappings $\{\rho'\ : R \rightarrow T'\}$.³ Moreover, it is logically impossible for a formal representation $\phi(R_i, R_j)$ to be false, although degrees of approximate morphism can be built into more elaborate conceptions of formal representation. In contrast, were notation ' $\phi(R_i, R_j)$ ' to denote a cognitive representation, $\phi(R_i, R_j)$ would be a pattern, not a fact; what it signified would be relative to o but not to an outside

observer's arbitrary choice of a mapping relation, and its depiction could well be erroneous. Confusion between the formal and cognitive senses of representation is one reason why information-processing generalities in modern psychology are seldom more than inchoate metaphors.

How well do the "neural representations" conjectured by Roitblat fit into this tripartite scheme? Despite his opening declaration that this label means nothing more than behavioral mediation, his subsequent statements read more like representation intuited as secondary aboutness. No other reading makes sense in view of his pervasive descriptions of mediators as representations specifically of the organism's salient environment rather than of the many other entities to which they are even more nearly isomorphic, and his unhesitant assumption that these code information for o about the world.

But if that is so, then Roitblat's claim that the research he cites demonstrates representations in animals is quite unwarranted. We still know almost nothing about the specific psychonomic mechanisms that embody even the most exemplary instances of cognitive aboutness, but surely mediators abound, especially in lower animals, that are neither primary cognitions nor vehicles for their elicitation. In all likelihood these often have a functional/compositional character intermediate between the cognitive/noncognitive extremes; but to savor the nuances of their multifaceted similarities and seek a behavior-theoretic counterpart of chemistry's periodic table, we must first take pains to appreciate their larger differences.

When we can plausibly explain some particular behavioral phenomenon in terms of a computationally effective mediation theory that does not include aboutness as part of the mechanism (see Fodor 1980) and avoids magical words like "code" and "information" unless these are clearly defined for this application, then it is time to consider whether mediational processes like these may be the psychonomic basis of common sense cognitive representations.

NOTES

1. No, Virginia, whatever propositions and other compound concepts may be, they surely do not much resemble the bundles of associations proposed by current semantic-network models. For that matter, neither do we have any satisfactory theories of the cognitive-aboutness relation (see Rozeboom 1979).

2. That is, given a logic in which notation ' $F(a_1, \dots, a_n)$ ' designates a pattern comprising elements a_1, \dots, a_n distinctively positioned by index within structural frame $F(_, \dots, _)$, we envision a piecemeal establishment of elicitational connections $\{a_i \rightarrow b_i\}$ and $\{F_j \rightarrow C_j\}$ in consequence of which $F_j(a_1, \dots, a_n)$ elicits $C_j(b_1, \dots, b_n)$ as distinct, for example, from $C_j(b'_1, \dots, b'_n)$ for permutation $\langle b'_1, \dots, b'_n \rangle$ of $\langle b_1, \dots, b_n \rangle$.

3. For every relational system $\langle R, \phi \rangle$ and one-one mapping ρ of R onto T' , $\langle R, \phi \rangle$ is isomorphic under ρ to the relational system $\langle T', \psi' \rangle$ in which ψ' is the image of ϕ in T' under ρ' .

Metatheory of animal behavior

Erwin M. Segal

Department of Psychology, State University of New York at Buffalo,
Buffalo, N.Y. 14226

From 1930 to 1960 many American psychologists knew what science was and how to do it. They took the basic characterization of their science from their view of physics and the positivistic empiricistic philosophers of science. Some of their tenets were: One must be exceedingly careful to avoid metaphysics at all cost; do not introduce any mentalistic terms; the only theoretical terms that one may introduce are those that are carefully defined by observables or by the experimenter's manipulations or operations (Segal & Lachman 1972). The major goal of this psychology was to predict and control behavior, or, more specif-

ically, given the stimulus, to predict the response; and given the response, to identify the stimulus that caused it (Watson 1914). It was a rigid framework, but also a safe one. Psychology was fighting for respectability as a science. Psychologists had to be tough, both to ward off the muddleheadedness of those who would explain behavior with souls and spirits, and to be accepted by their tough-minded scientific colleagues.

Most experimental psychologists stopped worrying about the "ghost in the machine" decades ago. We have gradually opened up the head and stuffed all kinds of former ghosts into it. We now unashamedly admit that the head contains such mentalistic entities as memories, images, beliefs, recognitions, rules, and the like. We can do this because we have no more doubts and worries about either our status as scientists, or about these mentalistic concepts being mistaken for ghosts or entelechies. Science (primarily by advances in biology and computation) has moved the center of common sense toward some sort of sophisticated physicalism. When we speak of a memory we never doubt that it has a physical instantiation, even if we can make no specific claims as to the detailed nature of that instantiation. Not only that, but we also assume that all of our readers likewise have no doubts that the memory has a physical instantiation. All of our journal articles are written with the presupposition of some kind of physicalism. The physicalism we believe in is much more sophisticated than that of Helmholtz and his nineteenth-century colleagues. We know that information and control can be represented in the patterning of elements onto some medium. We know that different patterns can cause many varied outcomes of size, shape, and behavior. We know that a controller does not have to supply the energy to the thing controlled; it simply has to inform it of the action to be taken. We also know that mentalistic ideas, such as categorization, logical processes, and decision making, can be physically realized. Newell (1980) gives a justification for using mentalistic concepts to communicate about physical systems.

Roitblat is entering into the fray somewhat belatedly. He presents representation of previous experience in part as if it were a relatively novel idea, or at least a controversial one, although he admits that it is foolish to consider the concept controversial. Indeed, the concept is no longer controversial at all. Volume 13 of *The Psychology of Learning and Motivation* "contained not one chapter with an S-R orientation" and was reviewed in *Contemporary Psychology* by a "card carrying mentalist" (Brewer 1981, p. 761). Even the Skinnerians have no metatheoretical disagreements with using mentalistic terms. Functional relations between inputs and outputs are seen by many of them simply to be more feasible. Representation as a concept is hardly new. It is perhaps the primary concept in the field of cognitive science, a field broad enough to encompass the research pursuits of many of the readers of this journal (e.g., Bobrow & Collins 1975).

Because the concept of representation is neither new nor controversial, perhaps the purpose of Roitblat's article is to carefully delimit it by adding boundary conditions so that someone identifying a representational system would have his potential characterizations constrained in some meaningful manner. But Roitblat does not succeed here either. He identifies five components of a representational system, but these components are so vague that investigators in the field would not know how to use them to constrain their investigations and perhaps would not wish to if they could.

Consider the "domain" of a representation. Generally, when behavioral research is done the investigator sets up a particular set of environmental conditions that the organism has to deal with, such as the conditions of a rat running around elevated paths and tables to get food (the Maier 1929 study). Maier, for example, could not delimit the domain of representation to the setup of tables, food, and paths because only by analyzing the data could he find out the form and content of the rat's cognitive representation. It seems that in Maier's study the room was

represented in a more holistic fashion. What was represented in the rat both in terms of domain and content was an empirical discovery, not a metatheoretical delimitation. Perhaps what Roitblat should have said is that one should not presuppose what the domain of the representation is, or to what the domain extends, without empirical justification. Only by investigation can we know what experiences are related and what the form and content of the representation is. Interestingly, characterizations of the representation are logically and methodologically independent of any specification of the details of the mode of instantiation in a medium.

As I see it, the value of the Roitblat paper is as an implicit argument for a cognitive comparative psychology. I was intrigued by the research findings reported – for example that rats seem to store information somewhat holistically, and that pigeons seem to store analogies to the second of two paired events. Researchers should investigate how organisms and species differ in the registration of information: What the patterns of the representation systems are; how they change over time; how they are affected by new inputs (contingent or not); and how they control the responses made by the organism. The metatheory is simply an attempt to characterize the changes in the organisms as they are affected by experience. The research goal is to show the evolution of mind.

Expectancy: The endogenous source of anticipatory activities, including "pseudoconditioned" responses

Patrick J. Sheafor

Department of Psychiatry, Case Western Reserve University, School of Medicine, Cleveland, Ohio 44106

Roitblat's metatheoretical approach to learning (representation) is a useful and most welcome attempt to help free us from the antitheoretical dogmas of radical behaviorism that still distort and constrict perceptions concerning relationships between conditioning processes (representations) and conditioning phenomena and, therefore, relationships between the behavioral and the brain sciences. Historically, "Pavlov's method" provided the first hope for radical behaviorism (Watson 1916). Accordingly, it should not be surprising that today it also provides its last refuge – at least in some quarters. Specifically, I wish to comment on a currently popular and influential line of research that seeks to elucidate neural mechanisms of Pavlovian conditioning based on radical behaviorist assumptions (Thompson 1976; 1980). Perhaps no other line of research better illustrates the importance of Roitblat's metatheoretical view of representation than this strictly operationist view of conditioning. First I will sketch the necessary historical background.

The crusading behaviorist, Watson, did not want to regard behavioral measures as indirect means of "looking into the minds of animals" (Griffin 1976). Indeed, he seized upon Pavlov's method primarily as a weapon to use in his quixotic battle against the method of introspection. Watson (1970) thought that in order to rid psychology of the soul (mind) and the subjective (i.e., "unscientific") method of introspection, he had to rid it of all central brain processes – what Pavlov called "higher nervous activities." So, for ideological reasons, Watson simply moved all central nervous system functions out into the peripheral nervous system where they could be directly ("objectively" or "scientifically") measured using Pavlov's method. Whereas physiologists and (nonideological, nonradical) behaviorists like Sechenov (1965), Pavlov (1927), Konorski (1967), Lashley (1929; 1950), Tolman (1932), Zener (1937), Grether (1938) and many others distinguished between brain systems of mind and behavior, knowledge and action, learning and performance, conditioning processes and conditioning phenomena,

expectancy and response, in Watson's radical behaviorist crusade these two systems were identified as being one and the same (Gallistel 1980). [See also BBS multiple book review of Gallistel's *The Organization of Action*, BBS 4 (4) 1981.]

Whereas Watson simply asserted the identity of mind and behavior, Hull (1937, p. 14) provided an apparently sophisticated and very influential theoretical deduction (Theorem 10) that Pavlovian or classical CRs (conditioned responses) are "the physical basis of expectation, of intention, of purpose, and of guiding ideas" (representations). Once classical CRs (observable elements) had thus been equated with cognitive capacities (conceptual elements), subjects in classical conditioning experiments could no longer "expect" (intend, purpose, or have guiding ideas) – all they could do theoretically was acquire classical CRs of the muscles, glands, and organs of the body: No CR – no expectation. Once the expectancy–CR distinction had been theoretically abolished, "the principle of expectancy" (Hilgard & Marquis 1940, p. 87) was no longer deemed applicable in classical conditioning experiments and, accordingly, the methodological "requirements of true conditioning" (p. 41) were "empirically determined" based on the assumption that radical behaviorism was correct.

The hope of this Watson-Hull-Spence behaviorist tradition was that classical CRs, as direct measures (the physical basis) of expectations and other cognitive functions, would provide the objective behavioral linkage between neurophysiology and psychology: On the one hand, classical CRs as the presumed "simplest form of associative learning" could only be reduced to neurophysiology; on the other hand, classical CRs were to provide the basis for a completely objective behaviorist account of all more complex instrumental conditioning and learning phenomena. These hopes provided the theoretical framework for the Hilgard and Marquis text that was to structure American perceptions of conditioning and learning for a generation (Kimbble 1961).

The hope that simple classical CRs would explain more complex instrumental learning phenomena was not supported by experimental results. Rescorla and Solomon (1967) concluded that "what concomitance we do observe between instrumental behavior and peripheral CRs is due to mediation by a common central state" (p. 170), and suggested that "it might very well turn out that instrumental responding is as sensitive, or perhaps even more sensitive a measure of the effects of Pavlovian conditioning procedures than are the traditionally measured conditioned visceral or motor reflexes themselves" (p. 178). Accordingly, based on compelling behavioral evidence, Rescorla (1967) rejected the methodological "requirements of true conditioning" (Hilgard & Marquis 1940) that had been "empirically determined" based on the erroneous assumption (Hull's Theorem 10 deduction) that classical CRs = expectations (and other cognitive functions). Systematically, these two watershed papers of 1967 marked the end of the 50-year reign of radical behaviorism in American psychology and, in Estes's (1975) apt phrase, "the rediscovery of Pavlov." Whereas Pavlov (1927) used a wide variety of conditioning procedures to investigate "higher nervous activities" (look into the minds and brains of animals), radical behaviorism, in conjunction with empiricist methodological doctrines of operationism and positivism, had reduced Pavlov's method simply to empirical investigations of classical CRs of muscles, glands, and organs of the body.

Nevertheless, some empirically inclined experimentalists still maintain that classical conditioning "involves the *direct* measurement of conditioned responses (CRs)" while instrumental "paradigms involve *indirect* measurement of purported CRs" (Scandrett & Gormezano 1980, p. 120; emphasis added); and also claim that classical conditioning simply "references the set of empirical findings arising from the application of the classical conditioning paradigm . . . provided specification of the paradigm is methodologically adequate" (Gormezano & Kehoe 1975, pp. 146–47). Such strictly operationist (pre-1967)

views of conditioning provide refuge today for the last remaining and futile hope of radical behaviorism, namely that Pavlovian or classical conditioning as the presumed "simplest form of associative learning" (Thompson 1976) – reducible only to neurophysiology – will provide ideal preparations or model systems with which to search for the engram (the neural basis of this simplest form of associative learning). Many experimentalists today find such a strict operationist view of conditioning appealing because: "it offers the behavioral biologist clear definitions of what is meant by the term learning, and a clear exposition of the appropriate controls to run. The behavioral biologist interested in understanding the brain mechanisms underlying this behavior finds it very useful to know precisely what controls to run" (Teyler, Baum & Patterson 1975, p. 71). As indicated above, however, the appropriate controls to run and the methodological "requirements of true conditioning" (Hilgard & Marquis 1940) were "empirically determined" on the basis of the erroneous assumption that radical behaviorism is correct and, accordingly, on the erroneous assumption that the principle of expectancy is not applicable in classical conditioning experiments (cf. Bolles 1972; Grossberg 1975; 1980; Kamin 1968; 1969; Rescorla & Wagner 1972).

The fact than an animal learns to expect events is revealed by the anticipatory character of its behavior or other activities (Sheafor & Rowland 1974). When an animal expects a biologically significant event (i.e., a US – unconditional stimulus) in a particular context (context-US association), it is biologically constrained to respond (act or react) in some ways (expectancy-potuated responses) but not in others (expectancy-inhibited responses). The animal's US-expectancy is the endogenous source of anticipatory responses. But what the animal, expecting a US, will actually do critically depends on the affordances (Gibson 1977) provided to the animal in the conditioned context or environment; environmental affordances constitute the exogenous sources of anticipatory responses. When a hungry pigeon, for example, is placed into a Skinner box in which it expects food, there is an increased probability, in the absence of the expected food, that it will approach and peck at a momentarily lighted response key, even though that affordance has never before been explicitly paired with food (Hitzing & Safar 1970). Similarly, when thirsty rabbits are placed into a classical conditioning chamber in which they expect an intraoral squirt of water, there is an increased probability that unpaired or probe test stimuli (lights or tones) will trigger CJM-PG (cyclical jaw movement-pattern generator) activity; the larger the expected squirt, the higher the probability that test stimuli will trigger CJM-PG outputs (Sheafor & Gormezano 1972); many other context-US-expectancy conditioning parameters systematically affect the probability of CJM-PG triggering by test stimuli (Sheafor 1975). These are just two examples – in the pigeon and rabbit – illustrating the principle that anticipatory responses are biologically determined by the ongoing interaction between an organism's needs and expectations (representations) and its environment; environmental affordances may disinhibit (Pavlov 1927, p. 115), elicit (Grether 1938), release (Hearst & Jenkins 1974), trigger (Sheafor 1975), or support expectancy-potuated responses. Many other examples – from man to the gastropod mollusc *Aplysia* – could be cited if space permitted.

Perhaps the only surprising thing about these anticipatory response phenomena is that, in the common wisdom of textbooks, they are referred to as "pseudoconditioned" responses and have erroneously been dismissed for over 40 years now as "a regular artifact in many types of learning experiments" (Uttal 1978). As Gould (1981, p. 26) has observed: "No medium is so conservative as the textbook; errors are copied from generation to generation and seem to gain support by sheer repetition," particularly so when such errors reflect an ideology implicitly shared by a scientific community (Gould 1978; Kamin 1974) or provide systematic support for the practice of a particular normal science paradigm (Kuhn 1970). These anticipatory re-

sponse phenomena (pseudoconditioning) have been systematically but erroneously dismissed as artifactual simply because they contradict the radical behaviorist assumptions upon which the empirically determined methodological requirements of true conditioning (Hilgard & Marquis 1940) are based, that is, the empirically specified classical conditioning paradigm (Gormezano 1972; Gormezano & Kehoe 1975).

Today this error is systematically perpetuated by advocates of strictly operationist views of conditioning who claim that "true" classical conditioning "involves the *direct* measurement of conditioned responses (CRs)" (Scandrett & Gormezano 1980, p. 120; emphasis added), and by experimentalists who believe that "true" CR engrams can be found by: (1) seeking out "an extremely well-behaved Pavlovian response which shows virtually no pseudoconditioning" (Thompson 1980, p. 181), and then (2) these "true" CRs can "be tracked through the nervous system from altered response output to constant stimulus input" (Thompson 1976, p. 211; cf. also Moore 1979).

Any realistic biological search for engrams must begin by rejecting radical behaviorist doctrines: We have no direct measures of simple associative learning processes today in any of the behavioral or brain sciences, not even in a relatively simple animal like *Aplysia* (Carew, Walters & Kandel 1981; Walters, Carew & Kandel 1981). To search for simple associative learning engrams (associative representations internally encoded in the organism's nervous system) is to search for the evolved neural circuitry and cellular mechanisms of expectancies (Grossberg 1980), the endogenous sources of anticipatory (preparatory) organismic activities, including anticipatory responses "disinhibited" by "any extra stimulus" (Pavlov 1927, p. 115), that is, pseudoconditioning phenomena. Currently popular and influential searches for true CR engrams are radically misled and misleading because they begin by seeking out "ideal" response systems which do not reveal these anticipatory response phenomena and are, therefore, relatively insensitive behavioral indices of the animal's expectancies. As Wasserman (1981, p. 24) has said: "Animal psychology has always sought out more effective ways of 'looking into the minds of animals' (Griffin 1976). It must continue to do so. . . . It is the *organism*, in conjunction with its stimulus surroundings, that selects the behaviors in which it engages as learning unfolds. The experimenter's task is to pick, by whatever means are at hand, those response systems that unequivocally divulge the subject's learning [representations]. It is precisely at this point in our current ignorance of how learning maps into performance that we stumble in our efforts to erect viable theories of learning and behavior."

Historicism, behaviorism, and the conceptual status of memory representations in animals

Charles P. Shimp

Department of Psychology, University of Utah, Salt Lake City, Utah 84112

Roitblat's ambitious paper may be broken down into two parts for commentary purposes – one that reviews contemporary empirical literature on animal memory and a second that evaluates historical and sociological correlates of the evolution of behaviorism. The first part is an admirable summary of the recent intellectual history of the problem of the animal mind and of a good portion of modern work on animal memory. Students and researchers new to the area should find this part especially useful, although the coverage is apparently not intended to be comprehensive. For example, many will note the omission of Wagner's recent work on the priming model (Wagner 1976).

The second part, dealing with the implications of animal memory for behaviorism, seems more problematical. The following comments focus on two items from a wide range of

possible topics for dispute: First, the extent to which Roitblat's concept of a memory representation is consistent with the historicist philosophy of science to which he appeals at the end of his paper, and second, his views on the nature and contemporary status of behaviorism.

A central feature of the historicist position is that parts of science heretofore considered largely or entirely objective, such as experimental methodology, observational data, and empirical facts, are in part subjective, arbitrary, context-dependent, and theory-laden (Hanson 1969; Kuhn 1970; Lakatos 1970; Rorty 1979). This view appears to be somewhat inconsistently applied in Roitblat's paper, with its implications applied to Watsonian reflexology but not to the author's own ideas about memory representations. At times, a representation is clearly described as theory-laden, but at other times, it is portrayed as existing objectively or having a "physical instantiation" independent of theory, and to be the kind of thing one could stumble over in a walk through a woods or find in his garage. According to the historicist position, theoretical concepts in science are not this kind of thing (Hanson 1969; Kuhn 1970).

Roitblat appears to find this theory-free and objective interpretation of a memory representation to be part of its justification: It is not metaphysical because of its assumed physical existence. One wonders what it is about the notion of a memory representation that makes it immune from the kind of historicist analysis that lets us see that reflexology, once also considered purely objective, is a theory-laden and context-dependent notion. Like Griffin (1976), Roitblat seems to treat the question of animal cognition as though it were a relatively simple empirical question, one that can be settled once and for all by listing a large number of experiments reputedly showing evidence of animal cognition. Both a historicist and a sophisticated behavioral position, however, would suggest that the question is really not so much an empirical one as a conceptual one about what it means to talk about animal (or human) cognition (Harzem & Miles 1978; Wittgenstein 1968). [See also "Cognition and Consciousness in Nonhuman Species" *BBS* 1 (4) 1978.]

There is no harm done to the legitimacy of the study of animal cognition if it is acknowledged that the evolving notion of a memory representation is a theory-dependent one. A memory representation is not the kind of thing the objective existence of which one proves or disproves. Thus, Roitblat concludes that certain experimental outcomes are evidence for the existence of a memory representation: "These results indicate that pigeons and monkeys can represent both the item and order information." An alternative way of looking at what the author is doing in such cases is to see that he is giving examples of what he means when he says that pigeons and monkeys can represent both the item and order information. The results then are not evidence for the existence of a representation. They are examples of what the author means by a memory representation (see Collingwood 1978 for several illuminating examples of this point).

Some contemporary behaviorists have little use for the historicist position and see little point in referring to a recent cognitive revolution; and some can be expected to object to claims such as Roitblat's that the idea of a memory representation is useful either to explain the results he reviews or to motivate interesting new methods (Branch 1977; Catania 1973). In some ways, both the author and his potential critics can be seen to agree. Nearly everyone can be expected to agree with the author that Watsonian reflexology, linear-mediating-chain explanations of memory, and the empiricist epistemology from which these are derived, are inadequate (Anderson & Bower 1973; Shimp 1976a; Tolman 1932). Disagreement is sure to follow, however, from Roitblat's description of behaviorism as necessarily committed to these discredited positions. It can be argued that the philosophical position of behaviorism is far more resilient and powerful than – and indeed quite different from – what he calls the "old behaviorist paradigm."

The philosophical position of behaviorism needs to be sharply

discriminated from the cluster of particular experimental methods, results, and theories that have accumulated under the behavioral label. One can return to the beginnings of behaviorism to find the following statement. "If you will grant the right to use consciousness in the same way that other natural scientists employ it – that is, without making consciousness a special object of observation – you have granted all that my thesis requires" (Watson 1913, p. 175). Such a view does not ban cognitive language in general or the notion of a memory representation in particular. Only some phenomenologists and humanistic psychologists argue that an acceptable meaning of a cognitive word, or an appropriate evaluative criterion for a psychological concept, is in terms of private mental experiences that are not publicly observable. Clearly this is not what Roitblat means by cognitive words and it is not what Tolman (1932) or others (Shimp 1976a; 1981) have meant.

It can be argued that "memory is no more or less objective than "behavior" (Shimp 1981). Both ordinarily refer only to certain contexts consisting of experiments, publicly observable events and statistics based thereon, and organizing theoretical frameworks (Shimp 1976a; 1981). Hardly anyone evaluates a psychologist's research program, even in this post-cognitive revolution era, in terms of his own private introspections. They do what Watson recommended: They use the ordinary evaluative procedures of the natural sciences. By Watson's criterion, there needs to be no objection from behaviorists to the notion of a "memory representation" (see also Harzem & Miles 1978).

In the beginning was the word

J. E. R. Staddon

Department of Psychology, Duke University, Durham, N.C. 27706

Publication of a paper in the softer sciences is often a perilous voyage among rocks and whirlpools defined by taboo terms. A few years ago, in the more behavioristically inclined journals, words like "adaptation," "memory" and "representation" meant instant shipwreck for the hopeful manuscript. On the other hand, frequent incantations of "operant," "occasioned by," "control" and (a perennial favorite) "response emission" carried talismanic power to ward off all harm and ensure speedy publication.

Things are, of course, much improved nowadays. Images, representations, nodes (but not, apparently, antinodes), codes, maps, and compasses are common currency – to the point that the casual reader might fancy himself perusing a spy novel rather than an account of birds pecking at lights or rats running in mazes. We are now quite free of the narrowness that so retarded the study of animal learning in years past. Naïve notions like "stimulus" and "response" still cause some alarm. Thoughtful reviewers occasionally feel misgivings about papers too liberally sprinkled with what might be termed "data-bound" terms. But for the most part, the cognitivization of America proceeds apace.

The real problem is, of course, one of human psychology, not the psychology of rats and pigeons. Francis Bacon, an early cognitive scientist, wrote "Few follow the things themselves, more the names of the things, and most the names of their masters." After a while, most of us begin to behave like the behaviorists' pigeons: We forget the concept and deal only with the label. Superficial work has a better chance of getting by, therefore, if at least it contains the right words. Conversely, a piece that lacks them, especially if it strikes off in a new direction, is likely to encounter a cool reception.

When we know little, we must coin terms of vague meaning. At first, they literally refer to nothing: We talk of the pigeon's "representation" of the delayed match-to-sample situation, yet we have almost no idea what we are talking about. This is not necessarily *bad* science, just *early* science. The terms label an

area of interest, a behavioral tendency, an imperfectly formed theory. Others may not find the questions interesting, or the nascent theory congenial. Because the terms are at first grounded in sand, they make easy targets. A lack of interest in a new approach can therefore always be justified by the highest critical standards.

Once new terms have been translated into testable and rigorous theory, criticism is muted; now the data can decide. But during that inevitable intermediate period, the theorist is using terms he can't define to explain phenomena he cannot understand.

All this is not to argue against criticism, but to remind us that criticism should go beyond the level of terminology. For example, a few years ago, there was a radical-behaviorist attempt to specify a permissible theoretical and empirical language for describing experimental results. If this effort had been directed at the precise meaning of common terms, then none could fault it – terms such as "stimulus control" and "response emission" deserve much more penetrating analysis than they have received. But the effort was more in the spirit of public health – a sort of mental hygiene where foreign terminological bacteria were to be firmly excluded – rather than a thoughtful attempt to understand meanings. It is hard to imagine a more futile exercise.

Despite these terrible examples, the cognitive approach will no doubt develop its own special talismans. A bias against affective variables and dynamic processes is already apparent. Animals may not be pieces of putty to be molded by reinforcement, but they are not emotionless, Spockian rule-followers either. It is one thing not to be straitjacketed by parsimony, but another to abandon it like a lapsed dieter and go on a binge in which entities multiply like rabbits. Not all abstract discussions of images and representations are evidence for subtle conceptual thinking; not a few read like the private language of a computerized Hare Krishna coven. One hopes that "node," "store" and "information" will not acquire the same power to suppress thought as did "control," "operant" and "reinforcement" in times past.

In this evenhanded spirit, let's look at some of the terms in Roitblat's useful and interesting article. "Representation" is rightly the central notion. Yet it loses meaning if it is applied as widely as Roitblat seems to require – to reflexes, circadian clocks, orientation by digger wasps, and feeding regulation by mantids. So used, "representation" becomes almost indistinguishable from "theory" and consequently loses its usefulness.

If the mechanism is simple, as in the mantid case, then the term representation is superfluous: Feeding is determined in negative feedback fashion by gut distention, and that is that. It is possible to term gut distention a "representation" of the animal's recent feeding history, but such a locution is metaphorical. In similar fashion, one might speak of the worn step as a representation of generations of marching feet, or Johnny Carson's gray hair as a record of his marital problems.

But if the mechanism is more complicated, then it is often possible to separate it into two parts, a part that maps in some way on the external world, together with a set of processes that act on the first part. Presumably this is what we really mean by "representation" – cf. Dennett (1978, pp. 90–108), a penetrating critique of Fodor's (1975) rather different concept. Even in this case, however, deeper understanding will reveal some more elementary structure: A representation is not an atomic unit. For example, there are computer programs that will do symbolic algebra; they might be said to represent the rules of algebra. Yet there is no single thing in the program corresponding to these rules. Instead, simple elements only distantly related to algebraic manipulations are strung together in such a way as to produce the proper transformations. So it will surely turn out for representations such as cognitive maps. As everyone knows (and many often forget), there is no little map in the

head. Formation of a map is surely a very complicated affair that involves search routines to assess the visual environment, processes that reconstruct spatial relations from this largely serial information, and additional processes that compute routes based upon the reconstruction. Still other processes, response rules, use these map processes to guide behavior according to past experience with rewards and punishments. The term "cognitive map" is therefore an interim concept that we hope will serve to bridge the gap between the level of comprehension in common intentional language, and the much more satisfactory account to be provided by a computational theory. [See BBS multiple book review of O'Keefe & Nadel's *The Hippocampus as a Cognitive Map*, BBS 2 (4) 1979.]

"Working" and "reference" memory are two other terms that deserve some analysis. They seem to be used sometimes in a purely procedural sense – working memory referring only to factors that remain valid reinforcement predictors only on a given trial, reference memory to features that are valid throughout an experiment. But at other times, they seem to refer to separate processes or "stores." (This kind of ambiguity is a familiar one, shared by the distinction between classical and instrumental conditioning.) There is little basis for a distinction at the process level. What is required is a workable memory model that incorporates both temporal and nontemporal information in a testable way (see Dale & Staddon 1981 for a beginning attempt). The working/reference dichotomy does not seem to have been very helpful in this respect.

Animal versus human minds

H. S. Terrace

Department of Psychology, Columbia University, New York, N.Y. 10027

Most psychologists and philosophers still accept Descartes's division of organisms into two groups: Humans, who have the capacity to think, and animals, who do not. According to this point of view, animal behavior, no matter how elaborate and complex, can always be reduced to some configuration of reflexes in which thought plays no role. It is, of course, true that this attitude presupposes a number of important extensions of Descartes's model of the reflex – in particular, different types of conditioned reflex. It is, however, equally true that the different models of the conditioned response that have been proposed by twentieth-century behaviorists have been regarded as automatic and thoughtless reactions to some immediate stimulus.

The success of stimulus-response (S-R) models of conditioned behavior has led most modern learning theorists to argue that nothing is gained by trying to explain an animal's behavior by appeal to its mental life and that the sole mission of a behaviorist is to isolate and to manipulate environmental variables that control behavior (e.g., Skinner 1950, 1974). At best, an appeal to unobservable mental events (or to neural events or to activities of a conceptual nervous system) detracts from the proper concern of a behaviorist: that is, prediction of behavior as a function of environmental variables.

Alternatives to the S-R model have hardly gone unnoticed. Almost 70 years ago, Hunter succinctly characterized the kind of evidence that would allow one to argue that an S-R model was not sufficient to explain some instance of animal behavior. "If comparative psychology is to postulate a representative fact, . . . it is necessary that the stimulus represented be absent at the moment of response. If it is not absent, the reaction may be stated in sensory-motor terms" (Hunter 1913, page 21). The few attempts (e.g., Tolman 1959) to develop alternatives to the S-R approach have met with little success. Purported instances of animal cognition were effectively neutralized by S-R theorists who postulated various types of covert mediators that seemed to satisfy Hunter's "sensory-motor" type of explanation. As a

result, behaviorists had little reason to concern themselves with what went on inside an animal's head.

This state of affairs has changed markedly during the past five years. A variety of studies of spatial memory (Olton 1978; 1979), delayed matching-to-sample (Roitblat 1980), sensitivity to monotonic patterns of food magnitudes (Hulse 1978), memory of the locus of key pecks in a sequence of pecks to different keys (Shimp 1976b; Shimp & Moffitt 1974), recall of photographs (Wright, Santiago, Urcuioli & Sands, 1982, in press), recognition of sequences of different stimuli (Weisman & DiFranco 1981b; Weisman, Wasserman, Dodd & Larew, 1980) and serial learning (Straub & Terrace 1982; Terrace 1982, in press) have created unprecedented difficulties for an S-R model of learned behavior. In each of these studies the behavior in question could not be explained by reference to an immediately available stimulus or to some covert S-R mediator. Accordingly, it appears as if Hunter's conditions for postulating a "representative fact" have finally been met. Because, as Roitblat notes in his target article, even an S-R model assumes that S-R relationships are represented in the nervous system, it is important to recognize that what Hunter and other psychologists mean by a "representative fact" is a representation of an S-S relationship.

In view of the dramatic examples of animal memory that Roitblat has marshalled to illustrate the concept of representation in animals, I was surprised that so much of his discussion of this problem included examples such as thermostats, the control of feeding in mantids by the volume of their guts, and circadian rhythms in plants. It is of course true that in some sense these and other examples Roitblat cites illustrate how various systems, biological and otherwise, represent features of their environments. Such facts, however, are hardly news. What is news is that animals can encode stimuli and relationships between stimuli that are not immediately present, and that these codes are not mediated by their behavior. In short, it appears as if psychologists have, for the first time, shown that animals can think.

The existence of S-S representations in animals poses an intriguing problem in that it provides a basis for studying complex forms of memory in nonverbal creatures. Theories of human memory (e.g., Anderson and Bower 1973) invariably postulate some sort of verbal or symbolic mediation. The ability of an organism to order events in its memory without verbal mediation should provide an invaluable biological benchmark for comparing human and animal memory.

Recognizing the need for S-S representations in accounting for animal behavior is, of course, but a modest step in discovering their nature. At this stage, I am not persuaded that information-processing models will prove that helpful.¹ My major reservation stems from the large investment of human information-processing models in verbal and other types of symbolic processes – processes that are unlikely candidates as mediators of animal memory. Another problem is that applying concepts of information processing, as expressed in Palmer's (1978) metatheory of representation, does little more than shift the level of analysis without increasing our powers of prediction. For example, Roitblat's observation that the medium of animal representation is "neural/cognitive" seems only to confuse neural and cognitive levels of explanation. Both cognitive and noncognitive processes are represented in the nervous system. Given our present state of knowledge, what distinguishes cognitive and noncognitive processes is not their neural mechanisms but the fact that, in the case of cognitive processing, the animal's behavior is being controlled by stimuli that the animal somehow generates for itself rather than by stimuli that are manipulated by the experimenter.

The one interesting aspect of Palmer's metatheory that Roitblat applies to the analysis of animal memory is the code the animal uses when storing important features of its environment. Roitblat's research on this problem provides an interesting example of the importance of specifying how, in a delayed

matching-to-sample task, a pigeon encodes the sample and how that code is used in selecting the appropriate choice (Roitblat 1980). The problem of how a stimulus is coded can, however, be readily stated without recourse to metaphors and concepts from human information-processing theory (e.g., Lawrence 1963) – metaphors and concepts whose relevance remains to be shown.

Now that the legitimacy of animal cognition as a field of inquiry has been demonstrated, I see little to be gained by approaching it as one would approach human cognition. Rather than prejudge the nature of animal cognition, I think it would be well to keep in mind Piaget's many demonstrations of the differences between a child's and an adult's mode of thought. I suspect that differences between animal and human thinking will prove to be at least as profound.

NOTES

1. It is of interest to note dissenting views concerning the usefulness of such models in explaining human cognitive phenomena. See, for example, discussions of this question by Neisser (1976) and Miller (1981).

Internal representations and indeterminacy: A skeptical view

William R. Uttal

*Institute for Social Research, University of Michigan, Ann Arbor, Mich.
48109*

Roitblat's thoughtful paper reminds us of some exceedingly important issues, among which by far the most important is surely the general question of the indeterminacy of internal representations. Anderson's (1978) assertion that there is a fundamental limit to what we can know about internal mechanisms and the antithetical views of those psychologists who disagree with him raise a critical theoretical issue. I would like to add to this dialogue some material from another domain – automata theory – that I believe has great relevance and that must take priority over the verbal and quasi-verbal arguments of cognitive psychologists. Given some leeway in definition, I believe formal proofs exist that should (but probably won't) resolve the psychological debate on indeterminacy.

There are, of course, some who may see one or another kind of escape hatch from mathematical proof in terms of irrelevance of domain or definition of entities. Indeed, over the past few years, whenever I have raised the possibility of indeterminacy among my fellow psychologists, I have frequently encountered a general unwillingness to even discuss "this philosophical matter of secondary interest to our empirical science." The matter is not secondary; for the very reason that it deals with the fundamental philosophical foundations of our science, it is quite primary. Practical arguments "to continue doing what we are able to do," or to use the principles of aesthetics, parsimony, or consistency, do not adequately contradict specific formal proofs.

There are common-sense arguments aplenty for the question of whether the internal structure (or coded internal representation) of a closed system can be uniquely determined from input-output relationships – but these arguments favor a negative answer. For example, there is no way of knowing what internal code gives rise to the number 8 when one puts the numbers 4 and 2 into a computer. The computational mechanism on the inside could have been encoded in a binary, decimal, or hexadecimal language. As another example, in a series of competent (i.e., noise-free) language translators, one can never determine whether any one of the interposed translators is speaking in any particular language. But this kind of informal argument is unsatisfactory, and though it may satisfy some of our intuitive predilections, it is not compelling in the same sense as a mathematical theorem.

In a classic paper from automata theory (which unfortunately

seems to have been totally forgotten in the context of psychology's debate over internal determinacy), E. F. Moore (1956) offers several theorems concerning devices called "finite state automata." He refers to these theorems collectively as "an analogue of the uncertainty principle" in physics. Moore's second theorem, particularly relevant to this discussion, states: "Given any machine S and any multiple experiment performed on S, there exist other machines experimentally indistinguishable from S for which the original experiment would have had the same outcome." After proving this theorem, Moore states (p. 138), "This result means that it will never be possible to perform experiments on a completely unknown machine which will suffice to identify it from among the class of all sequential machines." Although Moore does not so state, I believe that an argument can be made that the number of "other machines" may be very large. As the complexity of the machine grows, Moore's theorem suggests that both the number of possible alternative machines and the number of experiments that would be required to specify a unique internal mechanism would grow very large very fast. For a machine like the brain, both numbers would certainly approach a functional, if not a literal, kind of infinity.

Given that this theorem holds true for the class of machines described as finite automata, we must next ask if the brain can properly be considered a member of this class. Moore defines a finite automaton as having the following characteristics: a finite number of components, a finite number of internal states, a finite number of input symbols, and a finite number of output symbols. This does seem to describe the brain-mind complex. Therefore, Moore's second theorem would also be definitive in resolving the question of psychological indeterminacy. The burden of proving otherwise now rests upon those psychologists who do believe that alternative internal representations can be distinguished.

Roitblat's counterargument, that strict requirements of equivalence are required for the indeterminacy postulate to hold in psychology, is not convincing. His assertion (that the postulate would only apply in situations in which codal, informational, and computational complexity can be demonstrated) actually begs the question. The problem of establishing codal and computational equivalences differs somewhat from the problem of internal representation but it is just as difficult to solve. Roitblat's counterargument is, therefore, circular.

Other theorems proven by Moore concern the indeterminacy of previous states of the machine, given only the final state. It is impossible, he shows, to determine the initial state or any intermediate state of a finite automaton knowing only the device's final state. The analogy to psychological learning processes is obvious and again poses a serious challenge to theoretical psychologists interested in modelling the dynamics of behavioral plasticity.

If Moore's theorems are not only relevant, but also definitive (as I suspect they are), then psychology finds itself in a doubly difficult position with regard to reductive theory. Moore's theorems suggest that the unique specification of an internal representation of a "black box" is impossible if we cannot open it. Yet, even if we open up the black box, there may be no complete solution available. I have argued elsewhere (Uttal 1981) that the sheer complexity of the neural networks of the brain (the proper level at which one must seek the equivalents of mind) also precludes any realist's hope of neuroreductive analysis.

If these two limits (the indeterminate nature of a black box and the practical limits on the analysis of a very complex system) together pose a genuine and fundamental limit to reductionism in psychological theory, then history may look back upon some of our current analytic and reductionistic theories with amusement. Perhaps we should now examine this issue with less prejudice and more critical acumen. It may be necessary to make a much sharper distinction between mathematics used as a

descriptive model and mathematics used as a putative means of specifying unique internal mechanisms.

Mind the brain

Martha Wilson

Department of Psychology, University of Connecticut, Storrs, Conn. 06268

Roitblat's analysis of the concept of an internal representation is timely, particularly for those of us who are concerned with the physical instantiation of mental events. Hebb (1955) pointed out some time ago that experimental psychologists have often proceeded on the basis of implicit and outmoded assumptions about the workings of the brain. It is equally true that data from neurophysiology and neuropsychology have often been interpreted on the basis of outmoded assumptions about the workings of the mind. The target article provides a service in describing behavioral evidence for enduring states of mind that influence subsequent events, because these data pose a challenge to neuroscientists who may have conceived their task to be simpler than it actually is.

However, to view this recent shift of emphasis as a revolution may be overstating its importance. It is, of course, a cognitive tradition, started by the Gestalt psychologists and continued by Tolman (1932), to demonstrate the necessity for postulating cognitive structures in animals, attempting to win the ideological battle on difficult terrain so that they could then claim the rest of the territory. Because the history of psychology represents a continuing debate on whether the proper object of study is behavior or mind, this latest shift should perhaps be viewed less as a radical reorganization in thinking than as a return to earlier preoccupations. A longer view might characterize the present interest in cognitive processes as yet another turn of the spiral that we all hope is an ascending one.

Although behavioral evidence for the existence and nature of internal representations in animals is useful for its own sake, such animal data are particularly valuable because they suggest physiological experiments that may illuminate the locus and mechanism of internal representations. There are a number of studies that have examined brain structures involved in representations associated with spatial maps (O'Keefe & Nadel 1978), delayed matching-to-sample (e.g., Wilson, Kaufman, Zieler & Lieb 1972), and sequence discrimination (e.g., Cowey & Dewson 1972), among others, and such studies have contributed to our knowledge of brain organization. Some of the behavioral studies described in the target article also appear to have potential for further illuminating the neural substrates of behavior. For example, given Roitblat's (1980) finding that confusability of stimuli in the response set increases errors in a delayed matching-to-sample task to a greater extent than confusability of stimuli in the sample set, it should be possible to design recording and lesion studies that would be informative about the structures in which the inferred recoding of stimulus information takes place. That is, by exploiting designs used in behavioral research, our understanding of how the brain functions in behavior should be enhanced.

The notion that studies of behavioral processes may aid in forming hypotheses about brain function has not been as well-appreciated as the complementary argument that studies of the brain may have something to say about models of behavioral functioning. It is recognized that the much-studied patient H.M., who suffered from medial, temporal-lobe damage, has provided data that speak unequivocally to issues in memory research (Wickelgren 1968). And lesion studies of cortical function in monkeys have provided a wealth of evidence that "association cortex" does not associate in a manner demanded by stimulus-response models of behavior (Pribram 1958). Thus, the number and placement of boxes and arrows in information-

processing models can be and have been modified by the results of physiological manipulation. By this logic, the search for physical instantiation is not a luxury to which we might aspire when it is "more convenient" but a source of information in itself about the workings of the mind.

But, as noted above, it has been less obvious that models of how the brain is organized could benefit from hypotheses generated on the basis of behavioral evidence of internal representations. Representations in neurophysiology have generally been taken to mean topographic mappings of various sensory fields in various brain areas. But representations so defined do not, in and of themselves, provide knowledge of the external world – at least not in higher organisms in which temporal and spatial context is important in determining the effects of stimulation (Wilson 1978).

The "necessity" for postulating an internal spatial coordinate system that is continuously updated by current information has been stressed by MacKay (1973) among others, in order to explain how we know where our bodies and other objects are located in space. And Helson's (1964) adaptation-level theory implies a long-lasting but modifiable pattern of neural activity that determines the perceptual effect of sensory input. The reductionist answer to the fact that maps of spatial and feature aspects of stimulation do not provide an account of perceptual experience has been to look farther down the line for neurons that could serve such a function. The point of this commentary is to suggest another strategy. In line with the behavioral evidence discussed in the target article, it might be more fruitful to look for neural activity that reflects what is known about cognitive organization rather than activity that mirrors the physical values of stimuli and responses. Transcending S-R descriptions of brain as well as mind should increase our understanding of those events that underlie our knowledge of the external world and provide the basis for perception and action.

The heuristic value of representation

Thomas R. Zentall

Department of Psychology, University of Kentucky, Lexington, Ky. 40506

Behaviorism arose as a reaction against the subjective theories of the structuralists. The behaviorists (e.g., Watson 1970) proposed that behavior could be explained by considering only external events (i.e., stimulus, response, and outcome), with little or no invocation of internal processing of information. The major disadvantage of this approach has been the tendency to try to fit all behavior into the same external framework. In its theoretical form (e.g., Hull 1943) behaviorism has gotten lost in a complex maze of environmental factors presumed to be responsible for the variability of behavior. Where external factors have failed to account for the data, the theory has retreated into the organism by defining an increased number of hypothetical constructs (e.g., r_g , s_g), but behaviorists have continued to deny the involvement of active internal processing of information by animals. In order to maintain the generality of this theory, however, one must either accept a bad (or at least awkward) fit of the data to the theory, or one must restate the theory in sufficiently vague terms to accommodate the data.

If, as Roitblat does, one views an organism as an active processor of information, then one can explain data that the behaviorists have otherwise had difficulty accounting for – for example, natural concept learning (Herrnstein, Loveland & Cable 1976), abstract concept learning (Zentall, Edwards, Moore & Hogan 1981), and stimulus coding (Roitblat 1980). But Roitblat goes beyond an attempt to provide a better account of data that are poorly explained by simpler theories. Roitblat feels that a new approach to the study of behavior should be adopted,

one that focuses on "representation" as an animal's active means of processing information and organizing the environment. He proposes that the notion of representation can explain, not only data that are difficult for traditional behavioral stimulus-response (S-R) theory to explain, but also data that are readily accounted for by S-R theory.

Although representation theory presents a relatively fresh approach to animal research, and one that has great heuristic value, it may suffer from some of the same problems of overgeneralization as traditional S-R theory. For example, S-R theory has tried to explain delayed response behavior in terms of overt (or covert) stimulus-response chains, in spite of the fact that the stimuli and responses generally cannot be specified. Similarly, Roitblat proposes that overt S-R chains are more parsimoniously viewed as "external representations." If the term representation can be applied to both internal events and external behavior, then does it not lose some of its effectiveness as an explanatory concept? For example, the analogy between an animal's overt response chain and a human's use of a "note-pad," an activity that implies intentionality, is just the kind of extrapolation that created problems for the structuralists.

In fairness to Roitblat, theories are often overstated for the purpose of drawing attention to the weaknesses of alternative explanations. And Roitblat does temper his position by advocating representation as a metatheory, rather than an alternate theory. Furthermore, representation theory does have historical roots in early cognitive approaches to learning (e.g., Tolman 1932; Lawrence 1963), and this earlier work has already provided a body of data-based literature that would not have been collected had a more traditional (S-R) view of animal behavior been exclusively maintained. Furthermore, the adoption of Roitblat's cognitive view should encourage the use of methodology derived from human information-processing research, and thus should expand our treatment of animal behavior. Ultimately, such an approach applied to animal behavior should allow tests of cognitive theories that have been developed to account for human behavior. Roitblat's suggestion that representation can serve as a common rubric to account for virtually all behavioral change may be a bit premature, considering the relatively undeveloped data base that Roitblat has supplied. But given the potential value that such an information-processing perspective on animal behavior could have for stimulating new research, serious attention should be given to Roitblat's argument.

sensational metatheory could be applied to the analysis of animal behavior. Each reader can judge for himself the extent to which I have succeeded in reaching these goals. The commentators had their own ideas. Ultimately the virtue of the approach I described in the target article depends on its heuristic value. Although it seems to some, such as Zentall and perhaps Sheafor, that the approach has strong potential, only time will tell.

1. About behaviorism: Double-think

Branch, Catania, and Epstein all claim that I have misrepresented behaviorism. Modern behaviorism, they claim, bears no resemblance to the S-R associationism of its predecessor. Behaviorism, they say, does not involve covert changes within an organism that correspond to the overt changes that are observed; the mainstay of modern behaviorism is emitted – not elicited – behavior. Finally, I have omitted Skinner's analysis of private events, which, according to Epstein, are the behaviorist equivalent of representations.

From these criticisms one might conclude that the main objection of these behaviorists is that I am beating a dead horse. Modern behaviorism's only resemblance to that horse is its name. One might also conclude that modern behaviorism readily includes terms equivalent to representation, that if I had only read Skinner's analysis of private events, then I would have known that cognitive and behaviorist views are equivalent (or at least similar; Logue, Shimp). Apparently, these conclusions would be mistaken, however. At the very least, they are not the conclusions to which **Branch, Catania and Epstein** arrive. Furthermore, it seems to me that the conclusions they draw can only be reached through double-think – that is, they are internally inconsistent.

As I argued in the target article, behaviorists avoided the problems of representation by overtly ignoring or denying them while tacitly assuming their existence. "The mentalistic problem can be avoided by going directly to the prior physical causes while bypassing feelings or states of mind . . . If all linkages are lawful, nothing is lost by neglecting the supposed nonphysical link" (Skinner 1974, p. 13). Statements such as this appear to bear a close resemblance to statements of the form: stimulus S produces (or occasions) behavior R. Furthermore, **Catania** tells us that in Skinner's writing there are numerous examples of problem-solving that involve the creation of a sometimes private (i.e., within the behaver's skin) discriminative stimulus, which then occasions appropriate behavior. These private stimuli and the sometimes private responses, which may "occur on a scale so small that [they] cannot be detected by others" (Skinner 1974, p. 103), are equivalent to the covert stimuli and responses Hull used and reflect the tacit assumption of internal/external correspondence. Whether or not modern behaviorists would agree with this statement, their behavior does seem to support the parallel. In contrast, **Branch** argues that no mediators are necessary and that the entire problem of mediation dissolves into a pseudoproblem if one merely abandons a nineteenth-century billiard ball view of science and accepts the possibility of action at a temporal distance. If we viewed behavior the way the physicist views magnetism,

Author's Response

Representations and cognition

H. L. Roitblat

Department of Psychology, Columbia University, New York, N.Y. 10027

For the most part, the commentaries generated by my target article were quite interesting, and even contained a few surprises. Before I embark on a reply to my critics, it might be useful to outline some of my basic reasons for producing the target article. First, I wanted to develop the concept of representation as a metatheory that would be useful in the study of animal and human behavior. Second, I wanted to show that representations could be conceptualized as informational, physically instantiated entities independent of any assumptions about introspection, consciousness, or mysticism (cf. Jaynes and Rozeboom). Third, I wanted to illustrate how the repre-

he asserts, then there would be no need for representations as mediators of experience. Experience would be seen to exert its influence directly.

Branch's conceptualization of science is untenable. Even in the case of magnets (or any other area of physics) action at a *spatial* distance is assumed to be mediated by magnetic or other kinds of fields. It seems far more reasonable to interpret behaviorism as involving tacit assumptions of S-R-like mediators than to interpret it in the magical way that **Branch** proposes.

Branch indicates another way in which the behaviorist and cognitive paradigms differ: Cognitive approaches emphasize prediction and interpretation but behaviorist approaches emphasize prediction and control. By control **Branch** means, for example, knowing how to improve a rat's performance in a radial arm maze. I, for one, would be happy to sacrifice an emphasis on this type of control for an increased emphasis on interpretation (understanding).

Catania argues that part of the behaviorist objection to the use of representation stems from Skinner's association of mental terms with nonphysical events. How, Skinner asks, "can a mental event cause or be caused by a physical one" (Skinner 1974, p. 10)? This view may be an indictment of Cartesian dualism, but it is not relevant to modern materialist views of representation (cf. **Segal**). Furthermore, unless we accept **Branch's** doctrine of causation at a temporal distance, the same question can be applied to behaviorist systems, viz.: How does an organism's reinforcement history come to control later behavior? Cognitive approaches may be accused of leaving organisms lost in thought, but the behaviorist approach, unless it specifies the mechanism that mediates between experience and action, is guilty of leaving organisms lost in history.

Misunderstanding of the material basis of representation is probably what led **Catania** to rehash the homunculus problem in the form of a postulated inner copy. Events produce changes in organisms that experience them, and, unless one wants to deny the existence of those changes, the representation must be in the animal. **Catania's** assertion that they can be placed just as easily in the experimenter misses this point.

Similarly, if we look at the holes punched in a piano roll it is true that we will not find music. Nothing in the concept of representation, however, demands that the representation be the thing represented, or even that it be isomorphic with the thing represented. For example, during perception the unidimensional temporal variation comprising the pitch of a sound is coded in either an isomorphic temporal pattern of nerve impulses in the cochlear nerve (the cochlear microphonic) or, at higher frequencies, by particular neurons in the cochlear nerve determined by the location of their endings and by the particular structural properties of the cochlea (Green 1976). The resulting neural activity (especially at higher frequencies) is not sound, but it does represent sound. If the concept of representation were limited to the inner copy view **Catania** is criticizing, then he would be right to reject it.

To return to the analogy of the piano roll, it is clear that the roll does represent music that was played by someone at some time. That the roll bears little resemblance to either the music or the performer who first played it is

irrelevant. Yet, every time I push the pedals on that piano, the original performer does not come in, sit down, and play; nor is the music unaffected by things that happen to the roll, such as tears or punctures. Therefore, the roll clearly mediates the piano's "experience" affecting later performance.

Branch, **Catania**, and **Epstein** are apparently convinced that the use of representation is too seductive and will lure psychology away from its proper path, and that "When what a person does is attributed to what is going on inside him, investigation is brought to an end (Skinner 1974, pp. 17-8)." This may be true for some behaviorists, but it is a misrepresentation of cognitive science. All the questions that are important to the behaviorist regarding antecedents and controlling factors are also important to the cognitivist. The behaviorist's questions are simply a subset of the interesting and answerable questions, although they may have different emphases than those of the cognitivist and may be couched in a different language.

2. The reification of representations

Another misconception found in the commentaries of **Branch**, **Catania**, and **Epstein** concerns an issue also raised by a number of other commentators: The ontological status of representations. Representations have exactly the same ontological status as genes (or, as **Lachman & Lachman** note, as molecules in Brownian motion). Mendel's experiments indicated that the combining of particulate genes, as opposed to the mixing of fluids, was the appropriate way to conceptualize inheritance. Genes, though not "observed," were proposed as *things*; their status was, and remains, that of a theoretical construct.

Similarly, it is uninformative to explain an animal's structure or behavior as resulting from its possession of genes – all organisms are assumed to possess them. Instead, particular genes, selected as a result of particular environments, affect behavior through as yet unknown mechanisms. Also, geneticists do not limit their discussions of genes to the chemistry of DNA. They consider such topics as behavioral features, family trees, phenotypes, and the like. Thus, there is not just one level of analysis at which to discuss and interpret genetic functioning; nor are single units or easily identified, visible structures the only legitimate units for the description of neurophysiological structure and function.

Other parallels can be drawn between genes and representations, but these should suffice to make the point. **Branch**, **Catania**, **Logue**, and **Shimp** each failed to make such connections; nor did they distinguish between having representations (having genes) and utilizing a particular representation system (genome) based on a particular experience, code, or the like.

3. Indeterminacy and historicist epistemology

On the basis of a paper by Moore (1956), **Uttal** argues, correctly, that my equivalence rebuttal to Anderson's (1978) indeterminacy claim is inadequate because it says only that the two "machines" under investigation have produced identical results in all of the experiments so far conducted.

The most important point of Moore's paper for our purposes is his demonstration that no experiment of finite length can be sufficient to identify uniquely any particular finite state machine out of a class of all finite state machines. Other machines exist that could produce exactly the same outputs under the same conditions. Uttal concludes that the indeterminate nature of a black box and the overwhelming complexity of neural circuitry in the brain provide a genuine and fundamental limit to reductionist psychology.

Moore's theorem is very important to psychology, but not for the reasons Uttal suggests. Not only the brain, but also all scientific enterprises fit the definition of a finite state machine – having a finite (though large) number of states, inputs, and outputs – and they are all indeterminate in exactly the way described by Moore. No experiment is ever sufficient to identify (i.e. "prove") any theory from among all possible alternative theories (Kuhn 1970; Lakatos 1970; Watkins 1978). Anderson's and Uttal's arguments are misleading (as was my rebuttal) because they emphasize the wrong side of the problem. One cannot ask what number of machines are still consistent with a set of observations; the best one can do is to formulate bold conjectures and test them against one another and against data, thereby rejecting some classes of machines. The indeterminacy of theories is not a practical problem as Logue suggests, but a problem of ultimate knowability (see also Lachman & Lachman and Harnad 1982, p. 9).

Moore's theorem is an analog of "nonjustificationism" (Lakatos 1970; Weimer 1979), the historicist position mentioned by Shimp. I can do little more here than to present a sketch of the nonjustificational position, but its essential features are: (1) a denial of the existence of a nonproblematic, infallible data base; (2) a denial that scientific knowledge or theories (if they are at all different) can be proved; (3) a denial that individual theories can be falsified or repudiated with certainty (cf. Popper 1962) in part because of problems with establishing *ceteris paribus* assumptions; (4) the assertion that the proper unit of scientific endeavor is the research programme (Lakatos 1970), which consists of sets of eternally conjectural propositions; (5) the assertion that science consists of critical comparisons of research programmes in light of observation and fact; (6) that research programmes, not individual theories, are accepted if they are empirically and theoretically progressive, explaining as much as their less acceptable alternatives while simultaneously predicting hitherto unexpected facts; (7) that progress is measured by the degree to which a research programme's theories lead to the discovery of novel facts – that is, to the extent that they provide a positive heuristic (Lakatos 1970; see also Radnitzky & Andersson 1978; Cohen, Feyerabend & Wartofsky 1976).

Shimp's criticism that the historicist position is unevenly applied in the target article rests on something of a misconception about the nature of theory and observation within that position. As mentioned above, representations have the same ontological status as genes. One may indeed stumble over some genes on a stroll through a nucleus, but that does not make them any less theoretical or problematical. Both genes and representations are posited to have an existential reality (Lachman & Lachman).

It follows from this position that Shimp is right in his assertion that animal cognition is not an empirical matter to be settled with ultimate finality, but is instead a conceptual matter. Cognition, as I understand it, involves a research programme, or, in Kuhn's more extended sense, a paradigm or disciplinary matrix (Kuhn 1970). It is an approach to the conceptualization of behavior, not subject to ultimate justification or to falsification, but eternally subject to criticism. The data presented in the target article cannot "prove" a cognitive position, but are offered instead as evidence of the programme's progressive nature and its positive heuristic (Hulse, Logue, Sheafor, Terrace, Zentall). Whether this programme constitutes a revolutionary paradigm shift in Kuhn's sense or is simply a programme shift is largely irrelevant and depends on the criteria one applies in making that decision (Catania, Danto, Wilson).

4. Range of representations

A number of commentators were concerned with the broad range of phenomena I chose to include under the label of representation. For example, Gopnik, Jaynes, Logue, Staddon, Terrace, and Zentall all question whether it is appropriate to extend the concept of representation to include phenomena such as the factors controlling mantid feeding. There seem to be two reasons to make that extension. First, gut contents (and the like) do control behavior as a function of the organism's experience. They are no more or less physical than any other kind of representation. Not all representations need to be primarily cognitive, although even in the case of the mantid some cognitive component may be involved. Furthermore, in the absence of an adequate criterion for classifying processes as cognitive or noncognitive, restricting the use of representation to the former would be a bootstrap operation of dubious value based on ill-formed notions of cognition [See also Toates: "Homeostasis and Drinking" *BBS* 2 (1) 1979 and LeMagnen: "The Metabolic Basis of Dual Periodicity of Feeding in Rats" *BBS* 4 (4) 1981.]

Second, in order to exclude phenomena such as gut-regulated feeding from consideration as representations, we would have to define representations by exclusion. Learning has traditionally been defined in this manner (Kimble 1961) as a relatively permanent change in behavior potential excluding (by fiat) habituation, fatigue, and the like. One concomitant of this definition by exclusion has been the assumption that all forms of learning constitute examples of the same phenomenon. Such an assumption seemed perfectly reasonable to adherents of general learning theory, but it no longer appears tenable (see Johnston 1981; Sheafor). I suspect that part of the motivation for excluding these phenomena from consideration is the tacit assumption that anything which appears as easy to understand as gut contents could not possibly be what we mean by representation (Rozeboom).

Jaynes argues that the metaphor supplied by the notion of representation gives us a false feeling of understanding because it relies on fallacious intuitions regarding our own consciousness of consciousness. I think his argument is mistaken, however. Representations are not

based on introspection; nor are they, as behaviorists might claim, simply labels for ignorance, covering up how little we really know (Lachman & Lachman's weak meta-theoretical claim). Representations cannot be equated with either feature detectors or metaphorical images. They are functionally, not "ejectively," defined.

Rozeboom's argument is similar to **Jaynes's** in that it, too, is an attempt to restrict representation to a subset of what I advocated, defined by intuitions of consciousness. For an element of a causal chain to be considered legitimately as a representation in Rozeboom's sense, it must meet criteria of nontriviality and complexity of a particular (though unspecified) type. It must involve "aboutness," for which Rozeboom has no theory (he is relatively sure, however, that it has something to do with the content of mental acts such as believing, fearing, and so on and does not resemble bundles of associations). His entire argument rests on ill-defined criteria based on his vague and idiosyncratic intuitions of conscious cognition. It is hardly surprising, then, that the representational systems I described, with their functional orientation, do not meet his criteria (see also Harnad 1982).

Dretske and **Gopnik**, on the other hand, take the more reasonable view that representations can utilize two kinds of coding relations. Representations₁ are arbitrary, symbolic, and informational. Representations₂ are causal. These commentators agree that it is difficult to object to the use of representations₂ when there is a clear causal chain; we may not want to call those links "representations," but we can agree on their function as mediators. **Dretske**, however, continues this line to argue that causal representations₂ differ from representations₁, not only in terms of the potential arbitrariness of representations₁, but also in terms of their information-carrying status. He claims that representations₂ do not carry information about their antecedents, but this is clearly mistaken. One item carries information about another item or event if it identifies the other (reduces uncertainty), however imperfectly, from among its alternatives. The representation may not provide complete information (i.e. reduce the set of remaining alternatives to one) but, to the extent that it does contain information, it is a mapping between features of the two events. This mapping does not require the copying of features, or even isomorphism (i.e. the use of a cipher in which the code is a simple feature-for-feature substitution; cf. **Nelson**, **Rozeboom**). It requires only that some function (in the mathematical sense) or process exist (neither of which need be a homunculus), which can interpret features of the representation as standing for particular features of the represented object or event. Causal representations₂, are true representations because they preserve information about the antecedent event by virtue of being caused by it (i.e. a causal relation is information).

For example, in some code that might be used by a military organization, the symbol "B" could stand for any arbitrary number of features in the original message, meaning perhaps, "We will attack the bridge at the X river crossing at dawn tomorrow, flying in low from the north." The code "B" reduces uncertainty about the message. Anyone with the proper process (i.e. a code book and the ability to read it) could regenerate some of the features of that message (i.e. the battle plan) but not others (e.g., the type face used in the original message).

The latter information is not represented and is lost.

It is because of the potential complexity of coding relations that I think **Danto** and **Terrace** are mistaken in equating representations with associationism. For Danto, the representational position is merely an elaboration of old S-R models in which, he asserts, the stimulus already had a denotative role that is now attributed to the representation; using representations merely moves the "standing-for" relation one step into the animal. This view is mistaken (or at least misleading) for three reasons. First, in S-R conceptualizations the stimulus does not stand for anything. It is the input to a reflex-like chain that results in (emitted or elicited) behavior. Second, even if representations had the same semiotic relationship as stimuli, the proposition that they have their own real existence, as derived from processes internal and external to the organism, gives them an importance beyond their mere semiotic status (e.g. indicating why some features are represented, not others). Third, the data seem clear in indicating a higher level of internal complexity than is easily accommodated by focusing only on the semiotic relationship between stimuli and representations or by equating representations with associations between stimuli as Terrace proposes.

Dretske describes a hypothetical example in which I am distressed by a broken vase. Seeing the shards provides no information, he claims, about the causal antecedents of the breaking. This claim is clearly incorrect, however. Seeing a broken vase does provide at least the information that something collided with it, the approximate force of that collision, the vase's location prior to the collision, and so on. As any reader of Sherlock Holmes knows, a cigar ash can provide sufficient evidence to identify a murderer. To be sure, in order for the ash to provide information, auxiliary knowledge must be utilized. This merely highlights the notion that the information content of a representation depends on the processes embedded in the particular system using it (**Delius**, **Nelson**, **Rozeboom**, **Staddon**). It also reminds us that cognitive coding is unlikely to be truly arbitrary. As **Nadel** and **Delius** recall, the representational systems used by organisms evolved as a result of natural selection. They are functional.

In my house a broken vase is only likely to be caused by one person (assuming **Dretske** has not come to visit). I agree that this evidence is largely circumstantial, prone to error, and insufficient to convict, but a broken vase might provide sufficient evidence to indict. The appropriate conclusion, then, concerns neither whether causal remnants exist, nor whether they contain any information about the antecedent events, but rather, what information they do contain and how it is used (**Dretske**).

5. Representations and neurophysiology

I was somewhat surprised at the interpretation **Delius** and **Wilson** placed on my statement that as cognitive psychologists we usually avoid being overly concrete about the neural instantiation of cognitive representations. I intended that statement to indicate that cognitive psychologists do not typically draw circuit diagrams and the like in discussions of cognitive processing. It was not

intended as a call for ignorance, implying that neural structures, whether physically or mathematically described, are provisionally irrelevant, or that neurophysiological thinking should be delayed. There are three reasons, however, why cognitive psychologists do not typically operate at such a level. First, as Wilson points out, the emphasis in the study of neural functioning has been on the search for neural circuits that mirror stimulus-response connections. This is exactly the emphasis that I have argued we should avoid. Second, the level of function usually analyzed in neurophysiology is different from that common in cognitive psychology. Third, some neurophysiological functions (we do not yet know which) have no more to do with cognitive functions than the structure of ink has to do with literature. None of this is meant as a criticism of neurophysiology. I wholeheartedly support a more synergistic interaction between behavioral, functional, and neural levels of analysis, each level informing and being informed by the others.

Nadel and Delius emphasize the role that species-specific factors may play in the representations used by animals. Delius reminds us that an adequate theory of animal memory must consider its functioning in an evolutionary context. Different species, because of unique specializations and adaptations, can be expected to utilize different representational systems. There can be little disagreement with that conclusion. Nadel, on the other hand, tries to push the ecological functioning argument further by claiming that only certain kinds of tasks are acceptable as objects of study by virtue of their special relationship to the animal's ecology. It is a serious mistake, he claims, to allow laboratory tasks to define the kinds of representations one studies. Not every type of knowledge is likely to be the "main theme" of its own representational system. He argues, in essence, that there should be a principled set of constraints that prescribes appropriate and reasonable tasks to be studied and excludes all others. He implies that the only reasonable system to study (i.e., the only system likely to meet these unspecified constraints) is rats (or perhaps other creatures) in space.

The fundamental error in Nadel's argument is that no such principled set of constraints is justifiable or even desirable. It was precisely the failure to recognize the handicap imposed by the use of such a list that led behaviorism down the garden path. Scientific progress depends on theoretical pluralism and ultimately relies on the intuitions of its practitioners for determining its direction. Other phenomena besides rats in radial arm mazes may turn out to be the most illuminating. As Tolman recognized, our only guide in this matter is our own inclination (Tolman 1959).

6. Working versus reference memory

Nadel and others also objected to the distinction between working and reference memory. Marshall reminds us that there are numerous ways to divide up memory functioning: with versus without record, episodic versus semantic, locale versus taxon, autobiographical versus procedural, and explicit versus implicit. These distinctions are not mutually exclusive and are probably not exhaustive either. Their major function is to remind us

that memory is probably not a unitary function based on a single structure or mechanism (Gopnik), but involves a cohort of phenomena that have the common feature of cognitively mediating the effects of experience over temporal delays. The usefulness of any distinction will depend on its heuristic success and on the relevant data.

Working memory is similar to what James (1890) called primary memory, but is devoid of any assumptions of consciousness (cf. Jaynes, Rozeboom). Working memory is defined functionally or operationally and no ejective claims are necessary. For example, working memory is known to be involved in a situation when a subject has available a vocabulary of responses (usually learned), the selection of which depends on features of an event that is no longer present. The distinction between working and reference memory may not be clear in all cases (e.g., one trial passive avoidance), but failure to consider it in others leads to obvious contradictions and apparent paradoxes. For example, one might be forced to conclude that a pigeon knows how to perform delayed matching-to-sample (DMTS) with long sample presentation and short retention interval durations but "forgets" how to perform when short sample or long retention intervals are used, even in the same session. Invoking generalization decrement will not help because choice accuracy continues to improve if even longer sample durations are used (Roberts & Grant 1974). The same kind of confusion is shown by Nelson in his commentary.

A functional definition, contrary to Staddon's assertion, is not the same as ambiguity between processes and operations, but it may be silent about the mechanisms or structures implementing the different functions. Working and reference memory may involve different mental or neural structures or they may reflect different levels of activation of the same structure. Nevertheless, they are functionally distinct.

A number of studies indicate a dissociation between working and reference memory. For example, Olton and Papas (1979) found dissociation in rats performing in a 17-arm radial maze. During training, eight of the arms were baited at the start of each trial; the remaining nine arms were never baited. By the end of training, the rats were very accurate in avoiding any arm not containing food (i.e., both the never-baited and the previously-baited but already visited arms). The rats then received lesions to the fimbria fornix region. Following recovery, the rats were again tested with the same procedure. At first, they made errors with both kinds of arms; after some retraining they stopped making errors with the never-baited arms but still failed to base entries into the other arms on the occurrence or nonoccurrence of previous visits. Olton and Papas concluded that the fimbria fornix region provides connections to the hippocampus necessary to the functioning of working memory, but not of reference memory.

Nadel and MacDonald (1980), on the other hand, tested rats in two different mazes. In one seven-arm radial maze three of the arms, determined by their locations, were never baited. In another eight-arm radial maze three arms, determined by textured insets independent of location (i.e. their location changed from trial to trial), were never baited. In both mazes four arms were baited at the start of each trial. Following training, the rats received lesions to the dorsal hippocampus. Lesions

to this area, as opposed to the fimbria fornix, appeared to affect working and reference memory equally, but had less effect on rats in the maze where only texture was relevant, irrespective of location. Contrary to Nadel's claims, however, this experiment does not provide evidence against a working/reference memory distinction. At best, it suggests that the dorsal hippocampus is not involved in working or reference memory per se, but is involved in spatial memory functioning, although this conclusion may also be erroneous (Gray 1982; see *BBS* multiple book review, *BBS* 5 (3) 1982.).

Nadel also cites an experiment by Roberts and Dale (1981) that is purported to provide evidence against a working/reference memory distinction. Roberts and Dale tested their rats each day on a number of successive trials and reported proactive interference (PI) from one trial to the next (as described in the target article). On the basis of this finding they argue that rats do not reset working memory at the start of each trial (resetting here means completely erasing), but they do not argue that the working/reference distinction is irrelevant. They also suggest that

... as alleys are entered on the maze, temporal markers may be placed on the appropriate alleys in the memorial representation. These markers allow the animal to discriminate the order in which alleys were entered. However, the more closely in time alleys are entered, the more difficult it becomes to discriminate the order in which they were entered . . . PI arises primarily from a failure of temporal discrimination. (Roberts & Dale 1981, p. 279)

As Roberts and Dale recognize, however, there is a problem with the temporal discrimination hypothesis: Lengthening the intertrial interval from 60 to 240 seconds failed to attenuate the PI effect. In addition, temporal discrimination would also predict temporal patterning of errors relative to the time at which that arm was entered on the prior trial. It is not clear exactly what pattern one should expect because the PI effect consists of erroneously skipping arms that were visited on the prior trial but not yet visited on the current trial. In any event, Suzuki, Augerinos, and Black (1980) found that the choices made before and after transposition of orientational stimuli were independent of one another (i.e. there was no pattern). Temporal discrimination was not, therefore, functioning in that experiment.

The important feature of the model presented by Roberts and Dale is the proposal that the working memory system labels representations of arms entered with some kind of temporally (not spatially) varying cognitive flag and only imperfectly resets those flags at the end of a trial. A flag may spontaneously reset, but it does not spontaneously move from one arm to an adjacent one. Setting a flag is equivalent to adding it to a list, and so, as Roberts surmises, the list and flag hypotheses are identical. In the target article I did not conclude, as Roberts claims, that rats use lists instead of cognitive maps. Rather, I argued that rats use both – a cognitive map in reference memory and a list of visited locations (not stimuli) in working memory. This should not be interpreted as a claim that cognitive maps are little aerial photographs in the head or that lists must be stored in one part of the brain and maps in another.

Similarly, Roberts appears to stress a nonexistent dis-

agreement in the matter of preservation of order information. Such ordinal information is unnecessary, he claims (cf. Roberts & Dale 1981), if one simply assumes that rats forget memories of earlier arm entries (e.g., the flags spontaneously reset) before forgetting later entrances. Order information, however, is precisely that difference between the memories of the earlier and the later entries that allows them to be forgotten differentially.

Finally I am puzzled by Nadel's assertion that a brain could not have the capacity to store transiently any kind of information. What could he mean by this? Even the most elementary acquaintance with neuronal functioning provides an answer to that question. The "place" and "misplace" cells described by O'Keefe and Nadel (1978; 1979) and others, provide one example of a brain transiently coding information. Similarly, Bell and Hunter (1981) describe a hypothetical neural circuit that models working memory as a reverberating loop consisting of a set of neural channels converging on a recoding station that "corrects" the signal being carried by making each circuit agree with the majority of the signals received from the loop (cf. Gray 1982). This scheme, called "periodic majority recoding," is robust, efficient, and conservative – both of the number of individual neurons needed and of the information being carried. I have no particular confidence that these hypothetical memory organs will actually be found, but this system is physically realizable given what we know about neurons and their connections.

6.1. Alternatives to working memory

Any single system alternative to a dual working/reference memory system will have to utilize temporal coding if it is to account for performance on most of the tasks described in the target article. In such a memory system, each event results in a unique record, tagged for its time of occurrence. What appears to be working memory is simply the process that discriminates the temporal properties of the tags. This proposal is very similar to that of Roberts and Dale (1981) with one significant difference. Roberts and Dale argued that temporal markers are placed on the representation of the appropriate alleys as they are entered. That is, they viewed the markers as being temporary, created or placed on every trial, and distinct from the underlying representation, which they view as relatively permanent. These distinctions, at least, are essential.

Any memory system capable of accurate performance in tasks like these faces the double problem of classifying repeats of events as tokens of the same type while still being able to identify which event or events are currently relevant in choosing among the known alternatives. Temporal discrimination allows the latter but is difficult to use for the former because each repeat of an event is stored separately.

Roberts and Dale (1981) recognized that temporal discrimination, even in the more sophisticated form they proposed, cannot fully account for radial maze performance. Evidence favoring prospective codes from other tasks, such as delayed matching-to-sample (DMTS) and delayed sequence discrimination, also argues against a unitary memory system. (Further evidence, more strong-

ly inconsistent with temporal discrimination, is contained in a forthcoming paper by Roitblat and Scopatz that shows that proactive interference in DMTS does not consist specifically of competition with the sample presented on the previous trial – i.e. a failure of temporal discrimination – nor does the association between choices made on successive trials depend on the duration of the intertrial interval – i.e. competition among traces on the basis of “strength.”) A system that temporarily activates nodes in an appropriate associative network might be designed to produce prospective-like patterns of performance, but this system, too, would involve a stable representation (the associative network) and a labile one (activation of specific nodes). Hence it, too, is a dual memory system.

7. Delayed matching-to-sample and prospective codes

Roberts correctly describes a heretofore unrecognized property of the discriminant distance confusion measure I used (Roitblat 1980). The discrimination ratio will decline (for birds 51 and 62) if the error distribution remains constant as the total frequency of errors increases. This property would allow a code that was entirely orthogonal to either samples or comparison stimuli to produce patterns of discrimination ratios similar to those obtained. This suggests that the bias against a sample code is stronger than I had suspected. The prospective memory code hypothesis, of course, predicts that the pattern of errors changes as a function of the retention interval duration.

Examination of Table 3 in my earlier paper (Roitblat 1980) indicates that error distributions do not remain constant and that the declining discrimination ratio is not due simply to an artifact of the measure. Discrimination ratios equivalent to those reported, but calculated on the basis of percentage of total error rather than discriminant distance, should show a slope opposite that produced with the old measure if prospective codes are in use. They appear to do so, showing a change at least as robust as that obtained with the old discriminant distance measure (though I have not calculated the statistics). The values in that table can be converted to frequencies by multiplying the complement of the tabled value by 300, which was the total number of trials on which the two relevant samples were presented.

Roberts's second argument against prospective codes in pigeon delayed matching-to-sample rests on a mistaken claim that a prospective code requires a double transformation of the memory but a retrospective code requires only a single transformation. Both prospective and retrospective code models, however, require exactly the same transformations in exactly the same order. Both models require the identification of the sample and the generation of the correct response. They differ only in the assumed time at which the latter operation occurs – immediately for the prospective code model and at the time the comparison stimuli are presented for the retrospective code model.

The comparison between these models depends on the argument that more similar items are more likely to be confused (i.e. mistaken for one another) the longer they are held in working memory. That is, the memory repre-

sentation is assumed to vary along certain dimensions. Roberts recognizes only one dimension of variation, that of strength. His temporal discrimination hypothesis (Roberts & Dale 1981), for example, argues that rats place temporally changing markers on the representation of each arm of a radial maze as it is entered and that they make decisions about which arm to enter solely on the basis of this temporal variation. A spatial code or cognitive map hypothesis makes clear, however, that time is not the only possible dimension of variation. Spatial similarity (proximity) might also be a part of the working memory representation. That it does not appear to be so is irrelevant. Similarity among the sample/comparison stimuli in delayed matching-to-sample, in contrast, does appear to be part of the pigeon working memory representation. The finding that first choice errors contain information about the sample (Roitblat 1980) implies that the effects of retention interval or sample duration are not due simply to an increase or decrease in the systematicity of choices. Errors are differentially likely following different samples even when the samples are identical to the choices and all are about equally discriminable. Therefore, the 0° representation that is confused with the 12.5° representation does not have to be retrieved from reference memory; it is an intrinsic part of the dimensional variability of the memory system over time (See Roitblat, in press, for a model of such a process).

8. Representations of reinforcer sequences

Capaldi's main point is quite correct: It is probably not an issue whether rats use rules to code sequences of reinforcers, but rather, under what conditions do they do so, and under what conditions do they simply use associations. His conclusion, again correct, is that we don't know. How might we go about finding out?

There are really two questions involved in our analysis of serial learning. One question is more or less empirical, the other more or less theoretical. The empirical question revolves around the technique used to explore serial learning of reinforcer magnitudes. Two very similar techniques have been in use. Although some might claim they are identical (e.g. Capaldi, Verry & Davidson 1980a; Roberts), these claims cannot serve as a substitute for data. Even apparently trivial differences have been found on occasion to produce profound effects. The solution to this problem is simple: Use Capaldi's technique to produce Hulse's main finding (i.e. a difference between monotonic and nonmonotonic sequences). A demonstration of this sort is not useless replication, as Roberts claims, but is essential to assure that we are all talking about the same phenomenon. This problem of task identity is what I meant in the target article when I said that Capaldi's claims were groundless. I did not mean to extend that criticism to all theoretical claims. That would be an unfortunate misunderstanding.

The “theoretical” problem does not have a correspondingly simple solution. Capaldi and his associates have argued that a paired associate model is (a) simpler than a rule-based model and (b) adequate to account for the data presented by Hulse and his associates. These are not groundless claims to be taken lightly. In the target article I presented a simulation based on what I believe to be an

appropriate version of the association/generalization model. This simulation is inconsistent with the obtained data. Therefore, the model, as currently instantiated, is not adequate. Doing part of the job is not sufficient in this case (Roberts) because, in order to account for the running speed to the terminal element, the model must predict the same running speed to the 0 and to the 3-pellet element in the weakly monotonic condition. If we take the former prediction seriously, then we are compelled to take the latter prediction equally seriously. That prediction was falsified, and so this version of the model must be rejected. (Hulse's model is also more vague than Capaldi's. It is difficult, therefore, to subject it to the same standards of falsification. This is a drawback.) Suppose, however, that another version of the association/generalization model could be derived that would fit the data in question. What then of the parsimony argument?

Capaldi correctly points out that simply knowing that a series corresponds to a monotonically decreasing rule is not sufficient to predict when the 0-pellet element will occur. If Hulse's model provided a mechanism for predicting specific events, then its relative simplicity would vanish. The point of Hulse's model (and the whole basis for my claims about relative simplicity) is that it assumes that the rats do not predict specific events, only general trends: The next element will be less than the previous element. It is true that some information about the previous element is necessary to make any prediction about the next element, but neither the prediction nor the assessment of the previous element is assumed by Hulse to be an interval level judgment. Capaldi's model does require an interval level judgment.

Finally, Capaldi points out that rule-based and association/generalization-based accounts are not necessarily incompatible. Both clearly involve memory. Our task should be to investigate the form of representation used. The road to that end requires constant critical appraisal of alternatives.

9. Comparative cognition

In his commentary Hulse outlines the relevance of the target article to an earlier "prospectus" for comparative cognition (Hulse, Fowler & Honig 1978). Although I was also, in a sense, part of that prospectus (Riley & Roitblat 1978), Hulse's portrayal of comparative cognition is not quite appropriate to the intent of the target article. For Hulse, comparative cognition is comparative mainly with humans. His primary concern appears to be whether we can find functions in animals that are similar to those found in humans. If we can find such parallels then the feature cannot be specific to humans and might cause a restructuring of our thinking about this function. As an example of this sort of parallel Hulse cites work on language acquisition in nonhuman primates with the implication that, only if work on animal cognition can have a similar impact on thinking about human cognition, is it living up to its promise. I see two problems with this position. First, I doubt that Hulse's archetypal example of good comparative cognition has had anywhere near the impact he implies. How much has training chimpanzees in sign language really affected thinking about human cognition? Second, this approach casts animal cognition

as a poor relative of "true" cognition. To adopt such a position means that the study of animal cognition must always be a catch-up game.

The goal of the study of animal cognition is to understand cognitive processes. Procedures and conceptualizations developed in the animal lab are not simply imitations of "real cognitive science," but can provide powerful tools and ideas for the development of a truly comparative cognitive science (i.e. one grounded in evolution). The study of human cognition is, for the most part, too tightly bound to language use. Language is not equivalent to cognition. Organisms, including humans, have cognitive processes that do not involve language. Only prelinguistic humans and nonlinguistic animals can provide a means for studying these nonlinguistic cognitive processes in a way that is "uncontaminated" by the presence of linguistic ability. If my paper is to have any impact on the study of human cognition it will be at least partially based on this argument and on whatever I can contribute to a clearer analysis of representations in general.

Finally, I would like to reiterate one of Terrace's astute observations. There is no reason to believe that cognitive processes in animals must in any sense parallel cognitive processes in humans. The evidence is clear that humans differ greatly in their cognitive processes – as a function of age, if nothing else. We must be careful, then, in our study of animal cognition, not to allow conventional wisdom regarding the "facts" of human cognition to overly constrain our thinking on animal cognition. Similarly, the variety of cognitive processes we discover in animals can provide a fresh look at alternatives for the study of human cognition. The continued development of the study of animal cognition (and cognitive science in general) rests on its autochthonous development, not on its imitation of human cognition.

References

- Anderson, J. R. (1976) *Language, memory and thought*. Hillsdale, N.J.: Lawrence Erlbaum Associates. [RL]
- (1978) Arguments concerning representations for mental imagery. *Psychological Review* 85:249-77. [AWL, taHLR, WRU]
- Anderson, J. R. & Bower, G. H. (1973) *Human associative memory*. Washington, D.C.: Winston. [RL, CPS, HST]
- Beals, R., Krantz, D. H. & Tversky, A. (1968) Foundations of multidimensional scaling. *Psychological Review* 75:127-42. [taHLR]
- Beatty, W. W. & Shavalia, D. A. (1980) Rat spatial memory: Resistance to retroactive interference at long retention intervals. *Animal Learning and Behavior* 8:550-52. [taHLR]
- Bell, R. C. & Hunter, J. E. (1981) Periodic majority recoding and reverberatory memory. *Journal of Theoretical Biology* 88:323-32. [rHLR]
- Berkeley, G. (1709) *Essay towards a new theory of vision*. In: *The works of George Berkeley, Bishop of Cloyne*, ed. A. A. Luce & T. E. Jessup. London: Nelson. 1948-1957. [ACD]
- Bever, T. G., Straub, R. O., Terrace, H. S. & Townsend, D. J. (1980) The comparative study of serially integrated behavior in humans and animals. In: *The nature of thought: Essays in honor of D. O. Hebb*, ed. P. Jusczyk & R. Klein, pp. 51-93. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR]
- Bitgood, S. C., Segrave, K. & Jenkins, H. M. (1976) Verbal feedback and the feature-positive effect in children. *Journal of Experimental Child Psychology*, 21:249-55. [taHLR]
- Blough, D. S. (1959) Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior* 2:151-60. [RE, taHLR]

References/Roitblat: Representation in animal memory

- (1975) Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology. Animal Behavior Processes*, 1:3-21. [taHLR]
- Bobrow, D. G. (1975) Dimensions of representation. In: *Representation and understanding: Studies in cognitive science*, ed. D. G. Bobrow & A. Collins. New York: Academic Press. [taHLR]
- Bobrow, D. G. & Collins, A., eds. (1975) *Representation and understanding: Studies in cognitive science*. New York: Academic Press. [EMS]
- Bolles, R. C. (1972). Reinforcement, expectancy, and learning. *Psychological Review* 79: 394-409. [PJS]
- (1975) Learning, motivation, and cognition. In: *Handbook of learning and cognitive processes*, ed. W. K. Estes, pp. 249-80. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR]
- Branch, M. N. (1977) On the role of "memory" in the analysis of behavior. *Journal of the Experimental Analysis of Behavior* 28:171-79. [CPS]
- Brewer, W. F. (1981). Review of G. Bower, ed., *The psychology of learning and motivation: Advances in research and theory. Contemporary Psychology*, 26:761-62. [EMS]
- Bruner, J. S. (1969) Modalities of memory. In: *The pathology of memory*, ed. G. A. Talland & N. C. Waugh, pp. 253-59. New York: Academic Press. [JCM]
- Capaldi, E. J. (1966) Partial reinforcement: A hypothesis of sequential effects. *Psychological Review* 73:459-77. [EJC]
- Capaldi, E. J., Blitz, R. D. & Molina, P. (1979) Serial anticipation pattern learning in two-element and three-element series. *Bulletin of the Psychonomic Society* 14:22-4. [WAR]
- Capaldi, E. J. & Molina, P. (1979) Element discriminability as a determinant of serial pattern learning. *Animal Learning and Behavior* 7:318-22. [EJC, WAR]
- Capaldi, E. J., Nawrocki, T. M. & Verry D. R. (in press) Difficult serial anticipation pattern learning in rats: Rule encoding vs. memory. *Animal Learning and Behavior*. [EJC]
- Capaldi, E. J. & Verry D. R. (1981) Serial anticipation learning in rats: Memory for multiple hedonic events and their order. *Animal Learning and Behavior* 9:441-53. [EJC]
- Capaldi, E. J., Verry, D. R. & Davidson, T. L. (1980a) Memory, serial anticipation pattern learning and transfer in rats. *Animal Learning and Behavior* 8:575-85. [EJC, tarHLR, WAR]
- (1980b) Why rule encoding by animals in serial learning remains to be established. *Animal Learning and Behavior* 8:691-92. [WAR]
- Carew, T. J., Walters, E. T. & Kandel, E. R. (1981) Associative learning in *Aplysia*: Cellular correlates supporting a conditioned fear hypothesis. *Science* 211:501-4. [PJS]
- Carter, D. E. (1971) *Acquisition of a conditional discrimination: A comparison of matching-to-sample and symbolic matching*. Ph.D. thesis. New York: Columbia University. [taHLR]
- Carter, D. E. & Eckerman, D. A. (1975) Symbolic matching by pigeons: Rate of learning complex discriminations predicted from simple discriminations. *Science* 187:662-64. [taHLR]
- Carter, D. E. & Werner, T. J. (1978) Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior* 29:565-601. [taHLR]
- Catania, A. C. (1979) *Learning*. Englewood Cliffs, N.J.: Prentice-Hall. [ACC]
- Charnov, E. L. (1976a) Optimal foraging: The marginal value theorem. *Theoretical Population Biology* 9:129-36. [taHLR]
- (1976b) Optimal foraging: Attack strategy of a mantid. *American Naturalist* 110:141-51. [taHLR]
- Claparède, E. (1911) Recognition et moiïte. *Archives de Psychologie* 11:79-90. [JCM]
- Cohen, L. R., Looney, T. A., Brady, J. H. & Acuella, A. F. (1976) Differential sample response schedules in the acquisition of conditional discriminations by pigeons. *Journal of the Experimental Analysis of Behavior* 26:301-14. [taHLR]
- Cohen, R. S., Feyerabend, P. K. & Wartofsky, M. W. (1976) *Boston studies in the philosophy of science*, vol. 39: *Essays in memory of Imre Lakatos*. Dordrecht, Holland: D. Reidel. [rHLR]
- Collingwood, R. C. (1978) *An autobiography*. Oxford: Oxford University Press. [CPS]
- Conrad, R. (1964) Acoustic confusions in immediate memory. *British Journal of Psychology* 55:78-84. [taHLR, WAR]
- Cooper, L. A. (1975) Mental rotation of random two-dimensional shapes. *Cognitive Psychology* 7:20-43. [taHLR]
- Cooper, L. A. & Shepard, R. N. (1973) Chronometric studies of the rotation of mental images. In: *Visual information processing*, ed. W. G. Chase. New York: Academic Press. [taHLR]
- Cowey, A. & Dewson, J. H. (1972) Effects of unilateral ablation of superior temporal cortex on auditory sequence discrimination in *Macaca mulatta*. *Neuropsychologia* 10:279-89. [MW]
- Cumming, W. W. & Berryman, R. R. (1965) Stimulus generalization. In: *Stimulus generalization*, ed. D. I. Mostofsky. Stanford: Stanford University Press. [taHLR]
- Cumming, W. W., Berryman, R. & Cohen, L. R. (1965) Acquisition and transfer of zero-delay matching. *Psychological Reports* 17:435-45. [taHLR]
- Dale, R. H. I. & Staddon, J. E. R. (1981) A theory of spatial memory. Unpublished manuscript. [JERS]
- D'Amato, M. R. (1973) Delayed matching and short-term memory in monkeys. *The psychology of learning and motivation: Advances in research and theory* 7:227-69. [taHLR]
- D'Amato, M. R. & Cox, J. K. (1976) Delay of consequences and short-term memory in monkeys. In: *Processes in animal memory*, ed. D. L. Medin, W. A. Roberts & R. T. Davis, pp. 49-78. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR]
- D'Amato, M. R. & Worsham, R. W. (1972) Delayed matching in the capuchin monkey with brief sample durations. *Learning and Motivation* 3:304-12. [taHLR]
- Davis, M. (1958). *Computability and unsolvability*. New York: McGraw-Hill. [RJN]
- Dennett, D. C. (1978) *Brainstorms: Philosophical essays on mind and psychology*. Montgomery, Vt.: Bradford. [JERS]
- Devine, J. D., Burke, M. W. & Rohack, J. J. (1979) Stimulus similarity and order as factors in visual short-term memory in nonhuman primates. *Journal of Experimental Psychology: Animal Behavior Processes* 5:335-54. [taHLR]
- Eccles, J. C. (1981) The modular operation of the cerebral neocortex considered as the material basis of mental events. *Neuroscience* 6:1839-56. [JCM]
- Epstein, R. (1981) On pigeons and people: A preliminary look at the Columbian Simulation Project. *Behavior Analyst* 4:43-55. [RE]
- Epstein, R. & Skinner, B. F. (1981) The spontaneous use of memoranda by pigeons. *Behavior Analysis Letters* 1:241-46. [RE]
- Estes, W. K. (1972) An associative basis for coding and organization in memory. In: *Coding processes in human memory*, ed. A. W. Melton & E. Martin, pp. 161-90. Washington, D.C.: V.H. Winston & Sons. [taHLR]
- ed. (1975) *Handbook of learning and cognitive processes: Conditioning and behavior theory*, Vol. 2. Hillsdale, N.J.: Lawrence Erlbaum Associates. [PJS]
- (1976) Structural aspects of associative models for memory. In: *The structure of human memory*, ed. C. N. Cofer, pp. 31-53. San Francisco: W.H. Freeman. [taHLR]
- Feldman, C. F. & Toulmin, E. S. (1975) Theoretical problems in structural psychology. In: *Nebraska symposium on motivation*, ed. W. S. Arnold, pp. 409-75. Lincoln: Nebraska University Press. [RL]
- Flaherty, C. F. & Largen, J. (1975) Within subjects positive and negative contrast effects in rats. *Journal of Comparative and Physiological Psychology* 88:653-64. [taHLR]
- Fodor, J. A. (1965) Could meaning be an r_m ? *Journal of Verbal Learning and Verbal Behavior* 4:73-81. [JCM]
- (1975) *The language of thought*. New York: Crowell. [JERS]
- (1980) Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences* 3:63-72. [WWR]
- Freud, S. (1895/1966) Project for a scientific psychology. In: *The standard edition of the complete psychological works of Sigmund Freud*, vol. 1 (1886-1899), ed. J. Strachey, pp. 295-387. London: Hogarth Press. [JCM]
- Gaffan, D. (1977) Response coding in recall of colours by monkeys. *Quarterly Journal of Experimental Psychology* 29:597-605. [WAR]
- Gallistel, C. R. (1980) *The organization of action: A new synthesis*. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR, PJS]
- Gardner, H. (1975) *The shattered mind*. New York: Knopf. [JCM]
- Garner, W. R., Hake, H. W. & Eriksen, C. W. (1956) Operationism and the concept of perception. *Psychological Review* 63:149-59. [RL]
- Gibson, J. J. (1977) The theory of affordances. In: *Perceiving, acting and knowing: Towards an ecological psychology*, ed. R. Shaw & J. Bransford, pp. 67-82. Hillsdale, N.J.: Lawrence Erlbaum Associates. [PJS]
- Gormezano, I. (1972) Investigations of defense and reward conditioning in the rabbit. In: *Classical conditioning II: Current theory and research*, ed. A. H. Black & W. F. Prokasy, pp. 151-81. New York: Appleton-Century-Crofts. [PJS]
- Gormezano, I. & Kehoe, J. E. (1975) Classical conditioning: Some methodological-conceptual issues. In: *Handbook of learning and cognitive processes*, Vol. 2. *Basic learning and conditioning processes*, ed. W. K. Estes, pp. 143-79. Hillsdale, N.J.: Lawrence Erlbaum Associates. [PJS]

References/Roitblat: Representation in animal memory

- Gould, S. J. (1978) Morton's ranking of races by cranial capacity. *Science* 200:503-9. [PJS]
- (1981) Quaggas, coiled oysters, and flimsy facts. *Natural History* 90:16. [PJS]
- Gray, J. A. (1982) *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press. [rHLR]
- Green, D. M. (1976) *An introduction to hearing*. Hillsdale, N.J.: Lawrence Erlbaum Associates. [rHLR]
- Griffin, D. R. (1976) *The question of animal awareness: Evolutionary continuity of mental experience*. New York: Rockefeller University Press. [CPS, PJS]
- Grether, W. F. (1938) Pseudo-conditioning without paired stimulation encountered in attempted backward conditioning. *Journal of Comparative Psychology* 25:91-6. [PJS]
- Grossberg, S. (1975) A neural model of attention, reinforcement and discrimination learning. In: *International Review of Neurobiology*, Vol. 18, ed. C. C. Pfeiffer, pp. 263-327. New York: Academic Press. [PJS]
- (1980) How does a brain build a cognitive code? *Psychological Review* 87:1-51. [PJS]
- Hagg bloom, S. J. (1980) Reward sequence and reinforcement level as determinants of S-behavior in differential conditioning. *Animal Learning and Behavior* 8:424-28. [EJC]
- Halgren, C. R. (1974) Latent inhibition in rats: Associative or nonassociative? *Journal of Comparative and Physiological Psychology* 86:74-8. [taHLR]
- Hanson, N. R. (1969) *Perception and discovery: An introduction to scientific inquiry*. San Francisco: Freeman, Cooper. [CPS]
- Harnad, S. (1982) Neoconstructivism: A unifying theme for the cognitive sciences. In *Language, brain and mind*, ed. T. Simon & R. Sholes. Hillsdale, N.J.: Lawrence Erlbaum Associates. [rHLR]
- Harzem, P. and Miles, T. R. (1978) *Conceptual issues in operant psychology*. Chichester: Wiley. [CPS]
- Hayes-Roth, F. (1979) Distinguishing theories of representation: A critique of Anderson's "Arguments concerning mental imagery." *Psychological Review* 86:376-82. [taHLR]
- Hearst, E. (1978) Stimulus relationships and feature selection in learning and behavior. In: *Cognitive processes in animal behavior*, ed. S. H. Hulse, H. Fowler & W. K. Honig, pp. 51-88. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR]
- Hearst, E. & Jenkins, H. M. (1974). *Sign-tracking: The stimulus-reinforcer relation and directed action*. Austin, Texas: Psychonomic Society. [taHLR, PJS]
- Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row. [MW]
- Herman, L. M. & Gordon, J. A. (1974) Auditory delayed matching in the bottlenose dolphin. *Journal of the Experimental Analysis of Behavior* 21:19-26. [taHLR]
- Hebb, D. O. (1955) Drives and the C.N.S. (Conceptual Nervous System). *Psychological Review* 62:243-54. [MW]
- Herrnstein, R. J., Loveland, D. H. & Cable, C. (1976) Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes* 2:285-302. [taHLR, TRZ]
- Hilgard, E. R. & Marquis, D. G. (1940) *Conditioning and learning*. New York: Appleton-Century-Crofts. [PJS]
- Hitzing, E. W. & Safar, T. (1970) Auto-shaping: The conditions necessary for its development and maintenance. *Psychological Record*, 20:347-51. [PJS]
- Holton, G. (1975) On the role of themata in scientific thought. *Science* 188:328-34. [RL]
- Honig, W. K. (1978) Studies of working memory in the pigeon. In: *Cognitive processes in animal behavior*, ed. S. H. Hulse, H. Fowler & W. K. Honig, pp. 211-48. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR]
- (1982) Working memory and the temporal map. In: *Information processing animals: Memory mechanisms*, ed. N. E. Spear & R. R. Miller. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR]
- Honig, W. K. & Wasserman, E. A. (1981) Performance of pigeons on delayed simple and conditional discriminations under equivalent training procedures. *Learning and Motivation* 12:149-70. [taHLR]
- Horridge, G. A. (1962) Learning of leg position by the ventral nerve cord in headless insects. *Proceedings of the Royal Society Bulletin* 157:33-52. [JJ]
- Hull, C. L. (1937) Mind, mechanism, and adaptive behavior. *Psychological Review* 44:1-32. [PJS]
- (1943) *Principles of Behavior*. New York: Appleton. [AWL, taHLR, TRZ]
- Hulse, S. H. (1978) Cognitive structure and serial pattern learning by animals. In: *Cognitive processes in animal behavior*, ed. H. Fowler & W. K. Honig, pp. 311-40. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR, HST]
- (1980) The case of the missing rule: Memory for reward vs. formal structure in serial-pattern learning by rats. *Animal Learning and Behavior* 8:689-90. [EJC, WAR]
- Hulse, S. H. & Campbell, C. E. (1975) "Thinking ahead" in rat discrimination learning. *Animal Learning and Behavior* 3:305-11. [taHLR]
- Hulse, S. H. & Dorsky, N. P. (1977) Structural complexity as a determinant of serial pattern learning. *Learning and Motivation* 8:488-506. [EJC, taHLR]
- (1979) Serial pattern learning by rats: Transfer of a formally defined stimulus relationship and the significance of nonreinforcement. *Animal Learning and Behavior* 7:211-20. [taHLR, WAR]
- Hulse, S. H., Fowler, H. & Honig, W. K. (1978) *Cognitive processes in animal behavior*. Hillsdale, N.J.: Lawrence Erlbaum Associates. [SHH, rHLR]
- Hunter, W. S. (1913) The delayed reaction in animals and children. *Behavior Monographs* 2:1-86. [ACD, taHLR, HST]
- (1920) The temporal maze and kinaesthetic sensory processes in the white rat. *Psychobiology* 2:1-17. [taHLR]
- James, W. (1890) *The principles of psychology*. New York: H. Holt. [rHLR]
- Jarrard, L. E. & Moise, S. L. (1971) Short-term memory in the monkey. In: *Cognitive processes of nonhuman primates*, ed. L. E. Jarrard, pp. 1-24. New York: Academic Press. [taHLR]
- Jaynes, J. (1976) *The origin of consciousness in the breakdown of the bicameral mind*. Boston: Houghton Mifflin. [JJ]
- Johnston, T. D. (1981) Contrasting approaches to a theory of learning. *Behavioral and Brain Sciences* 4:125-39. [rHLR]
- Kamin, L. J. (1968) Attention-like processes in classical conditioning. In: *Miami symposium on the prediction of behavior: Aversive stimulation*, ed. M. R. Jones. Miami: University of Miami Press. [PJS]
- (1969) Predictability, surprise, attention, and conditioning. In: *Punishment and aversive behavior*, ed. R. Church and B. Campbell, pp. 279-98. New York: Appleton-Century-Crofts. [PJS]
- (1974) *The science and politics of IQ*. New York: John Wiley & Sons. [PJS]
- Kendler, H. H. (1952) What is learned? A theoretical blind alley. *Psychological Review* 59:269-77. [EJC]
- Kimble, G. A. (1961) *Hilgard and Marquis' conditioning and learning*, 2nd ed. New York: Appleton-Century-Crofts. [rHLR, PJS]
- Koch, S. (1976) More verbal behavior from Dr. Skinner. *Contemporary Psychology* 21:453-57. [RL]
- Kolers, P. A. and Perkins, D. N. (1975) Spatial and ordinal components of form and perception literacy. *Cognitive Psychology* 7:228-67. [RJN]
- Konorski, J. (1967) *Integrative activity of the brain*. Chicago: University of Chicago Press. [PJS]
- Kosslyn, S. M. (1981) The medium and the message in mental imagery: A theory. *Psychological Review* 88:46-66. [taHLR]
- Krechovsky, I. (1932) "Hypotheses" in rats. *Psychological Review* 39:516-32. [taHLR]
- Kuhn, T. S. (1970) *The structure of scientific revolutions*. 2nd ed. Chicago: University of Chicago Press. [MNB, tarHLR, CPS, PJS]
- Lachman, R., Lachman, J. L. & Butterfield, E. R. (1979) *Cognitive psychology and information processing*. Hillsdale, N.J.: Lawrence Erlbaum Associates. [RL]
- Lakatos, I. (1970) Falsification and the methodology of scientific research programs. In: *Criticism and the growth of knowledge*, ed. I. Lakatos & A. Musgrave, pp. 91-196. London: Cambridge University Press. [rHLR, CPS]
- Land, E. (1964) The retinex. *American Scientist* 52:247-64. [JJ]
- Lashley, K. S. (1929) *Brain mechanisms and intelligence*. Chicago: University of Chicago Press. [taHLR, PJS]
- (1950) In search of the engram. In: *Symposium of the Society for Experimental Biology*, vol. 4, pp. 454-483. New York: Plenum. [PJS]
- (1951) The problem of serial order in behavior. In: *Cerebral mechanisms in behavior*, ed. L. A. Jeffress, pp. 112-45. New York: Wiley. [SHH]
- Lawrence, D. H. (1963) The nature of a stimulus: Some relationships between learning and perception. In: *Psychology: A study of a science*, vol. 5, ed. S. Koch, pp. 179-212. New York: McGraw-Hill. [taHLR, HST, TRZ]
- Lettvin, J. Y., Maturana, H., McCulloch, W. S. & Pitts, W. H. (1959) What the frog's eye tells the frog's brain. *Proceedings of the Institute of Radio Engineers* 47:1940-951. [RJN]
- Logue, A. W. (in press a) The origins of behaviorism. In: *Points of view in the modern history of psychology*, ed. C. Buxton. New York: Academic Press. [AWL]
- (in press b) The growth of behaviorism. In: *Points of view in the modern history of psychology*, ed. C. Buxton. New York: Academic Press. [AWL]

References/Roitblat: Representation in animal memory

- MacKay, D. M. (1973) Visual stability and voluntary eye movements. In: *Central visual information*, vol. 7. *Handbook of sensory physiology*, ed. R. Jung, pp. 307-31. New York: Springer-Verlag. [MW]
- Maier, N. R. F. (1929) Reasoning in white rats. *Comparative Psychology Monographs* 6:1-93. [MNB, taHRL, EMS]
- Maki, W. S., Jr. & Leith, C. R. (1973) Shared attention in pigeons. *Journal of the Experimental Analysis of Behavior* 19:345-49. [taHRL]
- Maki, W. S. Jr., Brokofsky, S. & Berg, B. (1979) Spatial memory in rats: Resistance to retroactive interference. *Animal Learning and Behavior* 7:25-30. [taHRL]
- Maki, W. S., Gillund, G., Hauge, G. & Siders, W. A. (1977) Matching to sample after extinction of observing responses. *Journal of Experimental Psychology: Animal Behavior Processes* 3:285-96. [taHRL]
- Maki, W. J., Moe, J. C. & Bierley, C. M. (1977) Short-term memory for stimuli, responses and reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes* 3:156-77. [EJC]
- Marshall, J. C. (1979) Disorders of language and memory. In: *Applied problems in memory*, ed. M. Gruneberg & P. Morris, pp. 249-67. London: Academic Press. [JCM]
- Marx, M. H. & Hillix, W. A. (1979) *Systems and theories in psychology*. New York: McGraw-Hill. [taHRL]
- Mason, M. & Wilson, M. (1974) Temporal differentiation and recognition memory for visual stimuli in rhesus monkeys. *Journal of Experimental Psychology*, 103:383-90. [taHRL]
- Miller, G. A. (1981) Review of *Perspectives in cognitive science*, ed. by D. A. Norman, *Science* 214:57. [HST]
- Miller, G. A., Galanter, E. & Pribram, K. H. (1960) *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston. [AWL]
- Milner, B. (1970) Memory and the medial temporal regions of the brain. In: *Biology of memory*, ed. K. H. Pribram & D. E. Broadbent, pp. 29-50. New York: Academic Press. [JCM]
- Mishkin, M. & Delacour, J. (1975) An analysis of short-term visual memory in the monkey. *Journal of Experimental Psychology: Animal Behavior Processes* 1:326-34. [taHRL]
- Moore, E. F. (1956) Gedanken-experiments on sequential machines. In: *Automata studies*, ed. C. E. Shannon & J. McCarthy, pp. 129-53. Princeton: Princeton University Press. [rHRL, WRV]
- Moore, J. W. (1979) Brain processes and conditioning. In: *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski*, ed. A. Dickinson & R. A. Boakes, pp. 111-42. Hillsdale, N.J.: Lawrence Erlbaum Associates. [PJS]
- Murdock, B. B. (1976) Item and order information in short-term serial memory. *Journal of Experimental Psychology: General* 105:191-216. [taHRL]
- Nadel, L. and MacDonald, L. (1980) Hippocampus: Cognitive map or working memory? *Behavioral and Neural Biology* 29:405-09. [LN, rHRL]
- Neisser, U. (1976) *Cognition and reality*. San Francisco: W. H. Freeman. [HST]
- Nelson, R. J. (1975) On machine expectation. *Synthese* 39:129-39. [RJN]
- (1976) On mechanical recognition. *Philosophy of Science* 43:24-52. [RJN]
- Newell, A. (1980) Physical symbol systems. *Cognitive Science* 4:117-33. [EMS]
- Newell, A. & Simon, H. A. (1972) *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall. [RE]
- Norman, D. A., Rumelhart, D. E. & the LNR Research Group (1975) *Explorations in cognition*. San Francisco: Freeman. [RL]
- Oatley, K. (1978) *Perceptions and representations*. New York: Free Press. [taHRL]
- O'Keefe, J. & Nadel, L. (1978) *The hippocampus as a cognitive map*. Oxford: Clarendon Press. [JCM, rHRL, MW]
- (1979) Précis of O'Keefe & Nadel's *The hippocampus as a cognitive map*. *Behavioral and Brain Sciences* 2:487-94. [rHRL]
- Olton, D. S. (1978) Characteristics of spatial memory. In: *Cognitive processes in animal behavior*, ed. S. H. Hulse, H. Fowler & W. K. Honig, pp. 341-73. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHRL, HST]
- (1979) Mazes, maps and memory. *American Psychologist* 34:588-96. [HST]
- Olton, D. S. & Collison, C. (1979) Intramaze cues and "odor trails" fail to direct choice behavior on an elevated maze. *Animal Learning and Behavior* 7:221-23. [taHRL]
- Olton, D. S., Collison, C. & Werz, W. A. (1977) Spatial memory and radial arm maze performance by rats. *Learning and Motivation*, 8:289-314. [taHRL]
- Olton, D. S., Handelmann, G. E. & Walker, J. A. (1981) Spatial memory and food searching strategies. In: *Foraging behavior: Ecological, ethological, and psychological approaches*, ed. A. C. Kamil & T. D. Sargent, pp. 333-54. New York: Garland STPM. [taHRL]
- Olton, D. S. & Papas, B. C. (1979) Spatial memory and hippocampal function. *Neuropsychologia* 17:669-82. [rHRL]
- Olton, D. S. & Samuelson, R. J. (1976) Remembrance of places past: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes* 2:97-116. [taHRL]
- Paivio, A. (1975) Neomentalism. *Canadian Journal of Psychology* 29:263-91. [ACC]
- Palmer, S. E. (1978) Fundamental aspects of cognitive representation. In: *Cognition and categorization*, ed. E. Rosch & B. B. Lloyd, pp. 259-303. Hillsdale, N.J.: Erlbaum. [SHH, RL, taHRL, HST]
- Pavlov, I. P. (1927) *Conditioned reflexes*, G. V. Anrep, Trans. London: Oxford University Press. [PJS]
- Perdeck, A. C. (1958) Two types of orientation in migrating starlings *Sturnus vulgaris* and chaffinches *Fringilla coelebs* as revealed by displacement experiments. *Ardea* 46:1-37. [taHRL]
- Peterson, G. B. & Trapold, M. A. (1980) Effects of altering outcome expectancies on pigeons' delayed conditional discrimination performance. *Learning and Motivation* 11:267-88. [taHRL]
- Popper, K. P. (1962) *Conjectures and refutations*. New York: Harper and Row. [rHRL]
- Pribram, K. H. (1958) Neocortical function in behavior. In: *Biological and biochemical bases of behavior*, ed. H. F. Harlow & C. N. Woolsey, pp. 151-72. Madison: University of Wisconsin Press. [MW]
- Plyshyn, Z. W. (1978) Imagery and Artificial Intelligence. In: *Minnesota studies in the philosophy of science*, vol. 9, *Perception and cognition: Issues in the foundations of psychology*, ed. C. W. Savage, pp. 20-22. Minneapolis: University of Minnesota Press. [RJN]
- (1979) Validating computational models: A critique of Anderson's indeterminacy of representation claim. *Psychological Review* 86:383-94. [taHRL]
- (1980) Computation and cognition: Issues in the foundations of cognitive science. *Behavioral and Brain Sciences* 3:111-69. [taHRL]
- Rachlin, H. (1980) *Behaviorism in everyday life*. Englewood Cliffs, N.J.: Prentice-Hall. [AWL]
- Rachlin, H., Battalio, R., Kagel, J. & Green, L. (1981) Mazimization theory in behavioral psychology. *Behavioral and Brain Sciences* 4:371-417. [AWL]
- Radnitzky, G. & Andersson, G. (1978) *Boston studies in the philosophy of science*, vol. 58: *Progress and rationality in science*. Dordrecht, Holland: D. Reidel. [rHRL]
- Rescorla, R. A. (1967) Pavlovian conditioning and its proper control procedures. *Psychological Review* 74:71-80. [PJS]
- Rescorla, R. A. & Solomon, R. L. (1967) Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review* 74:151-82. [PJS]
- Rescorla, R. A. & Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning II: Current theory and research*, ed. A. H. Black & W. F. Prokasy, pp. 64-99. New York: Appleton-Century-Crofts. [PJS]
- Riley, D. A. (1968) *Discrimination learning*. Boston: Allyn & Bacon. [taHRL]
- Riley, D. A., Cook, R. G. & Lamb, M. R. (1981) A classification and analysis of short term retention codes in pigeons. *Psychology of Learning and Motivation: Advances in Research and Theory* 15. [taHRL]
- Riley, D. A. & Roitblat, H. L. (1978) Selective attention and related cognitive processes in pigeons. In: *Cognitive processes in animal behavior*, ed. S. H. Hulse, H. Fowler & W. K. Honig, Hillsdale, N.J.: Lawrence Erlbaum Associates. [rHRL]
- Roberts, W. A. (1972) Short-term memory in the pigeon: Effects of repetition and spacing. *Journal of Experimental Psychology* 94:74-83. [taHRL]
- Roberts, W. A. and Dale, R. H. I. (1981) Remembrance of places lasts: Proactive inhibition and patterns of choice in rat spatial memory. *Learning and Motivation* 12:261-81. [LN, tarHRL]
- Roberts, W. A. & Grant, D. S. (1974) Short term memory in the pigeon with presentation time precisely controlled. *Learning and Motivation* 5:393-408. [tarHRL]
- (1976) Studies of short-term memory in the pigeon using the delayed matching to sample procedure. In: *Processes of animal memory*, ed. D. L. Medin, W. A. Roberts & R. T. Davis, pp. 79-112. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHRL]
- Roberts, W. A. & Smythe, W. E. (1979) Memory lists for spatial events in the rat. *Learning and Motivation* 10:313-36. [taHRL]
- Roitblat, H. L. (1980) Codes and coding processes in pigeon short-term memory. *Animal Learning and Behavior* 8:341-51. [tarHRL, WAR, HST, MW, TRZ]
- (1982, in press) Discrete vs. graded processes in pigeon working memory.

References/Roitblat: Representation in animal memory

- In: *Harvard symposium on the quantitative analysis of behavior*, ed. M. L. Commons & A. R. Wagner. Cambridge, Mass.: Ballinger. [rHLR]
- Roitblat, H. L. & Scopatz, R. M. Sequential effects in delayed matching to sample. Submitted for publication. [rHLR]
- Roitblat, H. L., Tham, W. & Golub, L. (1982) Performance of *Betta splendens* in a radial arm maze. *Animal Learning and Behavior* 9. [taHLR]
- Rorty, R. (1979) Philosophy and the mirror of nature. Princeton: Princeton University Press. [CPS]
- Rozeboom, W. W. (1979) On behavioral theories of reference. *Philosophy of Science* 46:175–203. [WWR]
- Sands, S. F. & Wright, A. A. (1980) Serial probe recognition performance by a Rhesus monkey and a human with 10- and 20-item lists. *Journal of Experimental Psychology: Animal Behavioral Processes* 6:386–96. [taHLR]
- Scandrett, J. & Gormezano, I. (1980) Microprocessor control and A/D data acquisition in classical conditioning. *Behavior Research Methods and Instrumentation* 12:120–25. [PJS]
- Sechenov, I. M. (1965) *Reflexes of the brain: An attempt to establish the physiological basis of psychical processes*. Trans. S. Belsky. Cambridge, Mass.: MIT Press. [PJS]
- Segal, E. M. & Lachman, R. (1972) Complex behavior or higher mental process: Is there a paradigm shift. *American Psychologist*, 27:46–55. [taHLR, EMS]
- Shavalia, D. A., Dodge, A. A. & Beatty, W. W. (1981) Time-dependent effects of ECS on spatial memory in rats. *Behavioral and Neural Biology* 31:261–73. [taHLR]
- Sheafor, P. J. (1975) "Pseudoconditioned" jaw movements of the rabbit reflect associations conditioned to contextual background cues. *Journal of Experimental Psychology: Animal Behavior Processes* 104:245–60. [PJS]
- Sheafor, P. J. & Gormezano, I. (1972) Conditioning the rabbit's jaw-movement response: US magnitude effects on URs, CRs, and pseudo-CRs. *Journal of Comparative and Physiological Psychology* 81:449–56. [PJS]
- Sheafor, P. J. & Rowland, V. (1974) Dissociation of cortical steady potential shifts from mass action potentials in cats awaiting food rewards. *Physiological Psychology* 2:471–80. [PJS]
- Shimp, C. P. (1976a) Organization in memory and behavior. *Journal of the Experimental Analysis of Behavior* 26:113–30. [CPS]
- (1976b) Short-term memory in the pigeon: Relative recency. *Journal of the Experimental Analysis of Behavior* 25:55–61. [taHLR, HST]
- (1981) How objective is behavioral language? Talk given at a symposium at the 1981 annual meeting of the American Psychological Association, Los Angeles. [CPS]
- Shimp, C. P. & Moffitt, M. (1974) Short-term memory in the pigeon: Stimulus-response associations. *Journal of the Experimental Analysis of Behavior* 22:507–12. [HST]
- Simon, H. A. (1978) On the forms of mental representation. In: *Perception and cognition issues in the foundations of psychology*, ed. C. Wade Savage, pp. 3–18 Minneapolis: University of Minnesota Press. [taHLR]
- Skinner, B. F. (1945) The operational analysis of psychological terms. *Psychological Review* 52:270–77; 291–4. [RE]
- (1950) Are theories of learning necessary? *Psychological Review* 57:193–216. [HST]
- (1953) *Science and human behavior*. New York: Macmillan. [ACC, AWL]
- (1957) *Verbal behavior*. New York: Appleton-Century Crofts. [RE]
- (1966) An operant analysis of problem-solving. In: *Problem solving*, ed. B. Kleinmuntz. New York: Wiley. [ACC]
- (1969) Notes. In: *Contingencies of reinforcement: A theoretical analysis*, ed. B. F. Skinner, pp. 157–71. New York: Appleton-Century-Crofts. [ACC]
- (1974) *About behaviorism*. New York: Knopf. [rHLR, HST]
- (1977) Why I am not a cognitive psychologist. *Behaviorism* 5:1–10. [RE]
- Spence, K. W. (1937) The differential response of animals to stimuli differing within a single dimension. *Psychological Review* 44:430–44. [taHLR]
- Starr, A. & Phillips, L. (1970) Verbal and motor memory in the amnesia syndrome. *Neuropsychologia* 8:75–80. [JCM]
- Sternberg, S. (1975) Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology* 27:1–32. [taHLR]
- Stratton, G. M. (1927) *Theophrastus and the Greek physiological psychology before Aristotle*. New York: Macmillan. [ACC]
- Straub, R. O. (1979) *Serial learning and representation of a sequence in the pigeon*. Ph.D. thesis. New York: Columbia University. [taHLR]
- Straub, R. O., Seidenberg, M. S., Bever, T. G. & Terrace, H. S. (1979) Serial learning in the pigeon. *Journal of the Experimental Analysis of Behavior* 32:137–48. [taHLR]
- Straub, R. O. & Terrace, H. S. (1982) Generalization of serial learning in the pigeon. *Animal Learning and Behavior* 9:454–68. [AWL, taHLR, HST]
- Suzuki, S., Augerinos, G. & Black, A. H. (1980) Stimulus control of spatial behavior on the eight-arm maze in rats. *Learning and Motivation* 11:1–18. [tarHLR, WAR]
- Terrace, H. S. (1982, in press) Simultaneous chaining: The problem it poses for traditional chaining theory. In: *Harvard symposium on the quantitative analysis of behavior*, ed. M. L. Commons & A. R. Wagner. Cambridge, Mass.: Ballinger. [taHLR, HST]
- Terrace, H. S. & Straub, R. O. (1981) Generalization of serial learning in the pigeon. Paper presented at the Annual Meeting of the Eastern Psychological Association, New York, N.Y. [AWL]
- Teyler, T. J., Baum, W. M. & Patterson, M. M. (1975) Behavioral and biological issues in the learning paradigm. *Physiological Psychology* 3:65–72. [PJS]
- Thompson, R. F. (1976) The search for the engram. *American Psychologist*, 31:209–27. [PJS]
- (1980) The search for the engram, II. In: *Neural mechanisms in behavior*, ed. D. McFadden, pp. 172–222. New York: Springer-Verlag. [PJS]
- Thompson, C. R. & Church, R. M. (1980) An explanation of the language of a chimpanzee. *Science* 208:313–14. [SHH]
- Thorndike, E. L. (1911) *Animal intelligence*. New York: Macmillan. [taHLR]
- Thorpe, W. H. (1950) A note on detour behavior with *Ammophila pubescens* Curt. *Behaviour* 2:257–64. [taHLR]
- Tinbergen, N. & Kruyt, W. (1938) Über die Orientierung des Bienenwolfs (*Philaenus triangulum* Fabr.) I, II. Die Bevorzugung bestimmter Wegmarken. *Zeitschrift für vergleichende Physiologie* 25:292–334. [taHLR]
- Tolman, E. C. (1932) *Purposive behavior in animals and men*. New York: Century. [taHLR, CPS, PJS, MW, TRZ]
- (1948) Cognitive maps in rats and men. *Psychological Review* 55:189–208. [taHLR]
- (1959) Principles of purposive behavior. In: *Psychology: A study of science*, ed. S. Koch. New York: McGraw-Hill. [rHLR, HST]
- Tulving, E. (1972) Episodic and semantic memory. In: *Organization of memory*, ed. E. Tulving & W. Donaldson, pp. 381–403. New York: Academic Press. [JCM]
- Uttal, W. R. (1978) *The psychobiology of mind*. Hillsdale, N.J.: Lawrence Erlbaum Associates. [PJS]
- (1981) *A taxonomy of visual processes*. Hillsdale, N.J.: Lawrence Erlbaum and Associates. [WRV]
- Varela, F. J. & Maturana, H. R. (1981) Living ways of sense-making. Presented at the International Symposium "Disorder and Order," Stanford University, Palo Alto, California, September 13–15. [JJ]
- Von Uexküll, J. (1957) Stroll through the worlds of animals and man. In: *Instinctive behavior*, ed. C. H. Schiller, pp. 5–80 London: Methuen; orig. publ. 1934. [taHLR]
- Wagner, A. R. (1976) Priming in STM: An information-processing mechanism for self-generated or retrieval-generated depression in performance. In: *Habituation*, ed. T. J. Tighe and R. N. Leaton, pp. 95–128. Hillsdale, N.J.: Lawrence Erlbaum Associates. [CPS]
- Walker, J. A. & Olton, D. S. (1979) The role of response and reward in spatial memory. *Learning and Motivation* 10:73–84. [taHLR]
- Walters, E. T., Carew, T. J. & Kandel, E. R. (1981) Associative learning in *Aplysia*: Evidence for conditioned fear in an invertebrate. *Science* 211:504–6. [PJS]
- Wasserman, E. A. (1981) Response evocation in autoshaping: Contributions of cognitive and comparative-evolutionary analyses to an understanding of directed action. In: *Autoshaping and conditioning theory*, ed. C. M. Locurto, H. S. Terrace & J. Gibbon, pp. 21–54. New York: Academic Press. [PJS]
- Wasserman, E. A., Nelson, K. R. & Larew, M. B. (1980) Memory for sequences of stimuli and responses. *Journal of the Experimental Analysis of Behavior* 34:49–60. [taHLR]
- Watkins, J. (1978) The Popperian approach to scientific knowledge. In: *Progress and rationality in science*, ed. G. Radnitzky and G. Andersson. Dordrecht, Holland: D. Reidel. [rHLR]
- Watson, J. B. (1913) Psychology as the behaviorist views it. *Psychological Review* 20:158–77. [AWL, CPS]
- (1914) *Behavior: An introduction to comparative psychology*. New York: Holt, Rinehart and Winston. [EMS]
- (1916) The place of the conditioned-reflex in psychology. *Psychological Review* 23:89–116. [PJS]
- (1920) Is thinking merely the action of the language mechanism? *British Journal of Psychology* 11:86–104. [taHLR]
- (1970) *Behaviorism*. Reprinted. New York: W. W. Norton & Company. [AWL, PJS, TRZ]

References/Roitblat: Representation in animal memory

- Weimer, W. B. (1973) Psycholinguistics and Plato's paradoxes of the Meno. *American Psychologist* 28:15-33. [ACC]
- (1979) *Notes on the methodology of scientific research*. Hillsdale, N.J.: Lawrence Erlbaum Associates. [rHLR]
- Weimer, W. B. & Palermo, D. S. (1973) Paradigms and normal science in psychology. *Science Studies* 3:211-14. [taHLR]
- Weisman, R. G. & DiFranco, M. P. (1981a) Memory and recognition in delayed sequence discrimination. In: *Harvard Symposium in the Quantitative Analysis of Behavior*, ed. M. L. Commons & A. R. Wagner. Cambridge, Mass.: Ballinger. [taHLR]
- (1981b) Testing models of delayed sequence discrimination in pigeons: Delay intervals and stimulus durations. *Journal of Experimental Psychology: Animal Behavior Processes* 7:413-24. [HST]
- Weisman, R. G. & Dodd, P. W. D. (1979) The study of association: Methodology and basic phenomena. In: *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski*, ed. A. Dickinson & R. A. Boakes, pp. 337-63. Hillsdale, N.J.: Lawrence Erlbaum Associates. [taHLR]
- Weisman, R. G., Wasserman, E. A., Dodd, P. W. & Larew, M. B. (1980) Representation and retention of two-event sequences in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes* 6:312-25. [taHLR, HST]
- Wickelgren, W. A. (1968) Sparing of short-term memory in an amnesic patient: Implications for strength theory of memory. *Neuropsychologia* 6:235-44. [MW]
- Wilson, M. (1978). Visual system: Pulvinar-extrastriate cortex. In: *Handbook of behavioral neurobiology*, ed. R. B. Masterton, pp. 208-47. New York: Plenum Press. [MW]
- Wilson, M., Kaufman, H. M., Zieler, R. E. & Lieb, J. P. (1972) Visual identification and memory in monkeys with circumscribed inferotemporal lesions. *Journal of Comparative and Physiological Psychology* 78:173-83. [MW]
- Winograd, E. (1971) Some issues relating animal memory to human memory. In: *Animal memory*, ed. W. K. Honig & P. H. R. James. New York: Academic Press. [taHLR]
- Wittgenstein, L. (1968) *Philosophical investigations*. New York: Macmillan. [CPS]
- Worsham, R. W. (1975) Temporal discrimination factors in the delayed matching-to-sample task in monkeys. *Animal Learning and Behavior* 3:93-7. [taHLR]
- Wright, A. A., Santiago, H. C., Urcuoli, P. J. & Sands S. F. (1982, in press) Monkey and pigeon acquisition of same/different concept using pictorial stimuli. In: *Harvard Symposium in the Quantitative Analysis of Behavior*, ed. M. L. Commons & A. R. Wagner. Cambridge, Mass.: Ballinger. [tarHLR]
- Young, K. K. (1962) Tests of three hypotheses about the effective stimulus in serial learning. *Journal of Experimental Psychology* 63:307-13. [EJC]
- Zener, K. (1937) The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response. *American Journal of Psychology* 50:384-403. [PJS]
- Zentall, T. R., Edwards, C. A., Moore, B. S. & Hogan, D. E. (1981) Identity: The basis for both matching and oddity learning in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes* 7:70-86. [taHLR, TRZ]
- Zentall, T. R. & Hogan, D. E. (1978) Same/different concept learning in the pigeon: The effect of negative instances and prior adaptation to transfer stimuli. *Journal of the Experimental Analysis of Behavior* 30:177-86. [taHLR]