

# A "random-walk" simulation model of multiple-pattern learning in a radial-arm maze

IAN NEATH and E. J. CAPALDI  
Purdue University, West Lafayette, Indiana

Wathen and Roberts (1994) reported rather surprising results of a radial maze study that on any interpretation requires postulation of previously unsuspected high-level cognitive processes in rats. In each of four arms of the eight-arm radial maze, a different serial pattern unfolded over trials; for example, in one of the arms reward and nonreward alternated over successive trials. On each trial, rats came to track successfully four different patterns simultaneously. The authors suggested that rats tracked the pattern by using some form of trial-number strategy; that is, the trial number indicated which arms contained the better rewards. This strategy could involve a hypothesis, considered unlikely by some, that rats are capable of keeping track of as many as eight successive events—as, for example, by counting. A simulation model that embodies a specific form of the trial-number hypothesis is described here, and the results of the simulation correlate remarkably well with the observed data. In addition, the model makes four separate predictions that are supported by Wathen and Roberts's data and that seem beyond the scope of other available theories.

Wathen and Roberts (1994) reported an experiment in which rats were placed in an eight-arm radial maze and were tested for their ability to learn multiple patterns of reinforcement simultaneously. For example, if Arm *A* of the maze were assigned a single alternation pattern, there would be no food there on Trial 1, there would be food there on Trial 2, there would be no food on Trial 3, there would be food on Trial 4, and so on for eight trials. At the same time that the single alternation pattern was unfolding, three other different patterns of reinforcement were unfolding simultaneously in three other arms (see Table 1). Reinforced arms contained two 45-mg Noyes pellets, control arms contained one pellet, and nonreinforced arms contained no food. One pattern was single alternation (SA), beginning with nonreinforcement; a second was double alternation (DA), beginning with reinforcement; a third had two nonreinforced trials, four reinforced trials, and two nonreinforced trials (2N4R2N); and the fourth had the opposite, two reinforced trials, four nonreinforced trials, and two reinforced trials (2R4N2R). The remaining four arms were control arms; patterns were assigned randomly to the arms for each subject.

After 14 days of habituation to the maze, during which time the animals learned to push a food cover to obtain one pellet and to climb over a barrier that blocked entry to each arm, the training sessions began. The main behavioral measure was the mean rank of entry into each of the eight arms on each of the eight trials of each session. The main data of interest come from the 4 animals tested in Experiment 3, which were also tested on four extrapolation trials. On

these trials, the patterns were continued. Thus, the extrapolation trials for the single alternating pattern had no reinforcement on Trial 9, reinforcement on Trial 10, and so on. The extrapolation trials for the 2R4N2R had four nonreinforced trials. These data are reproduced in Figure 1 (closed circles).

There was multiple-pattern tracking for Trials 1–8, where mean rank of entry into the arms closely followed whether the arm choice was reinforced or not. That is, when there were two pellets in an arm, it was entered earlier than the control arms, which had just one pellet; when an arm had no food, it was entered later than the control arms. The important point to remember is that all four patterns were being tracked simultaneously. Remarkably, however, there appeared to be no evidence of pattern extrapolation on Trials 9–12, even for the most simple pattern, single alternation.

## THEORIES OF MULTIPLE PATTERN LEARNING

Theories of multiple pattern learning may be divided into two classes, the arm-pattern and the trial-number views. Both were originally designed to explain more simple pattern learning situations, but can be applied to multiple-pattern learning (see Wathen & Roberts, 1994). Within the arm-pattern class, one can further distinguish the sequential association view and the rule learning view. Both of these may be termed *horizontal* theories, because the focus is on the changing patterns of reinforcement within a particular arm over trials—or moving horizontally through Table 1.

### The Sequential Association View of Capaldi

According to this view (Capaldi, 1992; Capaldi & Verry, 1981), animals learn sequences of serial associations as chunks that are cued by prior events or memories. In an SA

Portions of this paper were presented at the Eighth Annual Indiana-Purdue-Kentucky Conference on Animal Learning and Behavior, West Lafayette, IN, March 1994. Correspondence may be addressed to I. Neath, 1364 Psychological Sciences Building, Purdue University, West Lafayette, IN 47907-1364 (e-mail: neath@psych.purdue.edu).

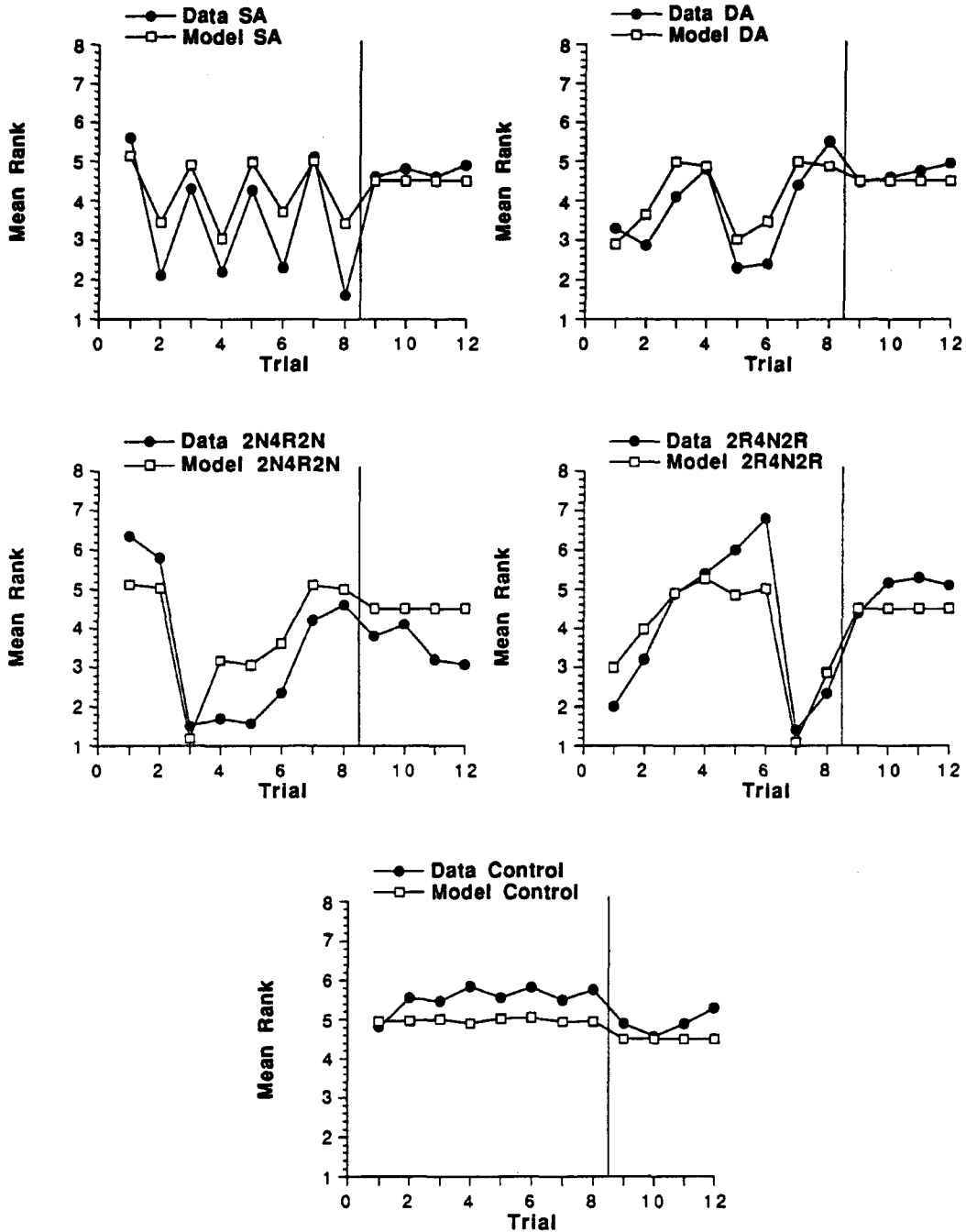


Figure 1. Mean rank of entry as a function of trial number for four patterns. The closed circles show data from Wathen and Roberts (1994, Experiment 3); the open squares show the results of the simulation.

pattern (NRNRNRNR), for example, memory of nonreward can signal an upcoming rewarded trial. In a DA pattern (RRNNRRNN), memory of a nonrewarded chunk can signal an upcoming rewarded chunk. In the case of multiple patterns, the animal would be required to keep track of the current trial so that the memory of reward (or nonreward) in a particular arm from the previous trial could signal the appropriate choice.

**The Rule Learning View of Fountain and Hulse**

According to this view (Fountain, 1990; Fountain, Henne, & Hulse, 1984), animals learn hierarchical rules for generating patterns by using both working memory and reference memory. Working memory is used to associate an event with a particular temporal/episodic context and is responsible for associating the sample stimulus with the trial. Reference memory, on the other hand, processes in-

Table 1

| Pattern | Trial Number |   |   |   |   |   |   |   |
|---------|--------------|---|---|---|---|---|---|---|
|         | 1            | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| SA      | N            | R | N | R | N | R | N | R |
| DA      | R            | R | N | N | R | R | N | N |
| 2N4R2N  | N            | N | R | R | R | R | N | N |
| 2R4N2R  | R            | R | N | N | N | N | R | R |
| Control | 1            | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Control | 1            | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Control | 1            | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Control | 1            | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Note—The patterns of reinforcement (R; two pellets) and nonreinforcement (N; zero pellets) for the four pattern arms are from Wathen and Roberts (1994). Control patterns always had one pellet. SA, single alternation; DA, double alternation.

formation independently of the context and is responsible for general rules and procedures. For multiple-pattern learning, reference memory is used to remember the overall sequence, such as *A*, then *B*, then *C*, and then *D*, whereas working memory remembers the immediately preceding response to mark position within the pattern (Olton, Shapiro, & Hulse, 1984). As with the sequential association view, chunking can occur.

According to both of these horizontal views, if animals are tracking patterns, they should correctly extrapolate the pattern, particularly the simple single and double alternation patterns, a prediction inconsistent with findings reported by Wathen and Roberts (1994). For successful tracking of multiple patterns, both views require that the animal remember at least two sets of hierarchical rules or sequential associations. For example, the rule-learning view requires the animal to remember an  $8 \times 8$  matrix of choices and to be able to identify correctly both the current trial and the appropriate arm choice from 64 possibilities. Although chunking the patterns might reduce the memory load, the matrix of response choices would still exceed reasonable estimates of memory capacity. Moreover, chunking would result in a matrix that is not square (e.g., a row of 8 for SA, a row of 4 for DA, and rows of 3 for 2N4R2N and 2R4N2R), which would complicate the interaction between reference and working memory. In contrast to these horizontal theories are a class of vertical theories that emphasize the patterns of reinforcement in all arms on a particular trial. These are “vertical” in the sense that the important comparison is the choices within a particular trial, or moving vertically through Table 1.

### The Trial-Number View

According to this view, rats keep track of the trial number, either by counting or by some other means (see, e.g., Holyoak & Patterson, 1981), and learn to expect reward in certain arms on each trial. This view requires that the animal is able to identify correctly the current trial from eight possibilities. Given correct identification of the trial number, the animal then expects certain arms to contain rewards. There is no requirement that the animal relate patterns of reinforcement on trial  $n$  to the patterns of reinforcement on trial  $n + 1$ . Data consistent with this view include sugges-

tions of numerical competence and counting ability, which would permit the trial number to be used as an important cue (see, e.g., Capaldi, 1993; Capaldi & Miller, 1988).

The trial-number hypothesis could be instantiated in many ways. The random-walk simulation model, described below, is one—perhaps the most simple—version of this view. According to the model, the animals in the multiple-pattern learning experiment do not learn horizontal patterns; rather, they learn vertical patterns that indicate which arm contains the maximum reward on each of the eight trials.

### THE RANDOM-WALK MODEL

The random-walk model assumes that a rat can correctly identify the current trial number. For all trials during the first session, all choices are random. As the animal experiences each reward level (zero, one, or two pellets), it remembers which arm contained the largest reward, and it associates that arm with the particular trial number. If two or more arms have the same large reward, the animal randomly associates the trial number with one of the arms.<sup>1</sup> On subsequent sessions, for each trial, the animal recalls which of the eight arms had the largest reward and enters that arm first. The remaining seven arms are entered in random order; hence the name of the model. Note that pattern learning consists of learning which one arm had the largest reward on each of eight trials and then randomly entering the remaining arms. The memory load is reduced from 64 items, according to the arm-pattern views, to 8 items. This reduction in memory load is achieved because of the assumption that the rat is capable of employing a complex cognitive solution, keeping track of eight successive trials.

For all simulations, there were 28 sessions, as in Wathen and Roberts's (1994) Experiment 3. For simplicity, it was assumed that there was no forgetting of trial number and that there was no forgetting of the largest reward magnitude, although these could be added as parameters. The simulation was run four times, and the data averaged to represent running 4 rats in an experiment and averaging over each subject's data. The patterns and reinforcement contingencies were those used by Wathen and Roberts (1994), and each virtual rat had a different random assignment of patterns to arms, as in the original. Note that there are no free parameters in the model: All of the parameters are specified in the study (e.g., number of arms, number of subjects, number of trials, and so forth). The results are shown in Figure 1 as open squares, along with the corresponding data from Wathen and Roberts (1994, Figure 7).

For Trials 1–8, the simulation is clearly replicating the observed data. Particularly interesting are Trial 3 of the 2N4R2N and Trial 7 of the 2R4N2R patterns, because on these trials, there is only one arm with the maximum reward. In the simulation, all 4 simulated rats will choose to enter this arm first, giving mean rank of entry of 1.0.<sup>2</sup> The rats in the Wathen and Roberts (1994) study showed an almost identical level of performance, with mean rank of entry of 1.5 and 1.4, respectively. Trials 1, 4, 5, and 8 had

two arms that had the maximum reward. The average rank of entry, in the simulation, for the arms with two pellets on these trials was 2.0. On Trials 2 and 6, there were three arms that had the maximum reward. The average rank of entry for these three arms was 2.5. The simulation model predicts that the more arms there are with the maximum reward, the later the mean rank of entry, exactly the pattern observable in the data. The reason, according to the model, is that the rats randomly associate one of the arms with the maximum reward as the arm to enter first. The other arms with a maximum reward will be entered later (on the average, fifth) and this will inflate the mean rank of entry.

The extrapolation trials were simulated by having the model randomly pick each arm to enter. Consider the following analogy. A human subject is presented with eight items in random order and is asked to recall all eight. On an extrapolation trial, the subject is now asked to recall the ninth item. If pressed for a response, the subject would have to guess (randomly pick a response), because the question does not make sense. The simulation assumes that the animals are recalling the trial number, but when this number is not valid, the animals randomly pick which arms to enter. Again, the simulation results mimic the observed data.

The correlations between the observed and predicted data were .958 for SA, .868 for DA, .855 for 2N4R2N, and .917 for 2R4N2R for the first eight trials, and they were .922, .858, .822, and .916 when the extrapolation trials were also included. The fits are by no means perfect, but they are certainly suggestive. A vertical model with no free parameters, which assumes only that the rat can recall the arm with the largest reward on each trial, produces results that closely mimic horizontal multiple-pattern learning. In addition, it predicts that there will be no extrapolation of the pattern, even for the most simple and regular of patterns, because horizontal pattern learning did not occur. Again, this is what was observed.

## IMPLICATIONS AND PREDICTIONS

In 1972, Underwood posed a question to researchers studying human behavior: "Are we overloading memory?" The same question may be asked of multiple-pattern learning theorists. Both horizontal views previously discussed can account for pattern learning when only a single pattern is tracked. However, when multiple patterns are tracked, these views require the animal to remember a substantial amount of information; in the case of the Wathen and Roberts (1994) study, an  $8 \times 8$  matrix. Moreover, both horizontal views predict accurate extrapolation, particularly for the relatively simple single and double alternation patterns. No evidence of extrapolation was found.

The vertical view offered here, the random-walk simulation model, predicts a failure to extrapolate. This model assumes that the animal has to remember only 8, rather than 64, items. Interestingly, even if one considers it questionable to assume that a rat can correctly identify the current trial number, note that like the vertical view presented

here, both horizontal views previously mentioned have the same requirement in order to explain the Wathen and Roberts (1994) data. The difference between the horizontal and vertical views is that any horizontal view must necessarily make an additional assumption of accurate recollection of eight times more information: In addition to tracking trial number, the animal must also track the information contained in the columns of Table 1. Thus, the trial number assumption is indispensable for any model.

The random-walk model is easily tested, in principle, because it makes at least four strong predictions that are supported by the Wathen and Roberts (1994) data. First, as previously indicated, the model predicts no extrapolation, regardless of the pattern, unless the subject believes that the first extrapolation trial is actually Trial 1 again. There is no extrapolation because the subject is using the trial number as a cue, and when the trial number is invalid, there can be no memory to guide the choice of an arm. Second, the model predicts highly stereotyped pattern of arm choice. Because the animals are learning which arm contains the maximum reward on that trial, they will always enter that arm very early, producing the very strong preferences for the initial arm choice of a trial observed by Wathen and Roberts. Third, the model predicts that mastery of a horizontal pattern is unrelated to the regularity of the pattern. For example, a random pattern could be learned better than a single alternation pattern because it is the vertical patterns that are more important than the horizontal patterns. Thus, in the Wathen and Roberts study, the rats "learned" the irregular 2N4R2N pattern as well as they learned the more regular DA pattern. Fourth, the model predicts less irregularity in performance when, for all trials, there is an equal number of arms that contain the maximum reward. For example, the two earliest ranks of entry were observed when there was only one arm that contained the maximum reward. When the maximum reward was available in three arms, the mean rank of entry was later.

It is important to note that the random-walk model is not applicable when single patterns are to be tracked: The model applies only when the task places a high load on the cognitive abilities of the organism. According to the model, when this situation occurs, the animal copes by reverting to a mnemonically simpler but cognitively complex strategy of associating trial number or position with the maximum reward available on that trial. Although most of the decisions (56 out of 64) will be random, the remaining 8 nonrandom decisions will give rise to pattern learning.

One interesting question to explore concerns how high the cognitive load must be before the animal uses the random-walk strategy. Cognitive load, however, should not necessarily be interpreted as a capacity limitation based on the number of items. In the human literature, for example, many theories of working memory are moving away from item- or time-based limits of capacity and instead focus on resource or discrimination limits (e.g., Nairne, in press). It is likely that capacity limitations in the rat, particularly in situations of high cognitive load, do not depend on the absolute number of items or the absolute passage of

time, but rather on the animal's ability to discriminate or differentiate between two or more alternatives on the basis of cues available at retrieval (cf. Neath & Knodler, 1994).

Historically, two approaches to serial learning—in both humans and animals—have been dominant competitors (cf. Crowder, 1976). The item–item view holds that prior items in a series provide cues for later items in the series. The position–item view holds that cues associated with either the absolute or relative position of an item in a series provide cues for items in the series. The trial-number view advanced here is obviously a version of the position–item view, but it differs from previous versions: The cue for an item identified here is not position in a series of items but rather position in a series of series.

#### REFERENCES

- CAPALDI, E. J. (1992). Levels of organized behavior in rats. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 385-404). Hillsdale, NJ: Erlbaum.
- CAPALDI, E. J. (1993). Animal number abilities: Implications for a hierarchical approach to instrumental learning. In S. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence* (pp. 191-209). Hillsdale, NJ: Erlbaum.
- CAPALDI, E. J., & MILLER, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 3-17.
- CAPALDI, E. J., & VERRY, D. R. (1981). Serial order anticipation learning in rats: Memory for multiple hedonic events and their order. *Animal Learning & Behavior*, **9**, 441-453.
- CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- FOUNTAIN, S. B. (1990). Rule abstraction, item memory, and chunking in rat serial-pattern tracking. *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 96-105.
- FOUNTAIN, S. B., HENNE, D. R., & HULSE, S. H. (1984). Phrasing cues and hierarchical organization in serial pattern learning by rats. *Journal of Experimental Psychology: Animal Behavior Processes*, **10**, 30-45.
- HOLYOAK, K. J., & PATTERSON, K. K. (1981). A positional discriminability model of linear-order judgments. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1283-1302.
- NAIRNE, J. S. (in press). Short-term/working memory. In E. L. Bjork & R. A. Bjork (Eds.), *Handbook of perception and cognition: Vol. 10. Memory*. New York: Academic Press.
- NEATH, I., & KNOEDLER, A. J. (1994). Distinctiveness and serial position effects in recognition and sentence processing. *Journal of Memory & Language*, **33**, 776-795.
- OLTON, D. S., SHAPIRO, M. L., & HULSE, S. H. (1984). Working memory and serial patterns. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 171-182). Hillsdale, NJ: Erlbaum.
- UNDERWOOD, B. J. (1972). Are we overloading memory? In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 1-23). New York: Wiley.
- WATHEN, C. N., & ROBERTS, W. A. (1994). Multiple-pattern learning by rats on an eight-arm radial maze. *Animal Learning & Behavior*, **22**, 155-164.

#### NOTES

1. This need not imply that a particular rat necessarily remembers the two (or more) alternatives and picks one at random. Rather, it means only that the basis of the choice is random. For example, if both Arm *D* and Arm *G* contain the maximum reward for a particular trial, *n*, Rat A might enter Arm *D* first and associate Arm *D* with trial *n* whereas Rat B might enter Arm *G* first and associate Arm *G* with trial *n*.
2. The actual values were 1.13 and 1.12, respectively, because on the first trial all arms are entered randomly.

(Manuscript received February 10, 1995;  
revision accepted for publication May 13, 1995.)