Journal of Experimental Psychology:
Animal Behavior Processes
2008, Vol. 34, No. 3, 415–418

0097-7403/08/$12.00 DOI: 10.1037/0097-7403.34.3.415

BRIEF REPORTS

Learning About Environmental Geometry: A Flaw in Miller and Shettleworth's (2007) Operant Model

Michael R. W. Dawson
University of Alberta

Debbie M. Kelly
University of Saskatchewan

Marcia L. Spetch and Brian Dupuis
University of Alberta

Many studies have examined how humans and other animals reestablish a sense of direction following disorientation in enclosed environments. Results showing that geometric shape of an enclosure is typically encoded, sometimes to the exclusion of featural cues, have led to suggestions that geometry might be encoded in a dedicated geometric module. Recently, Miller and Shettleworth (2007) proposed that the reorientation task be viewed as an operant task and they presented an associative operant model that appears to account for many empirical findings from reorientation studies. In this paper we show that, although Miller and Shettleworth's insights into the operant nature of the reorientation task may be sound, their mathematical model has a serious flaw. We present simulations to illustrate the implications of the flaw. We also propose that the output of a simple neural network, the perceptron, can be used to conduct operant learning within the reorientation task and can solve the problem in Miller and Shettleworth's model.

Keywords: reorientation task, geometric cues, feature cues, artificial neural networks, operant choice

The reorientation task is a paradigm that has been used extensively to study the types of information used by humans and animals to navigate in their environment (Cheng & Newcombe, 2005). In this task, participants are reinforced for going to a particular location in an arena that is typically rectangular in shape. The participant then has to find that location again after being disoriented, and possibly after changes have been made to the arena. This task is used to determine the geometric and featural cues that can be used to "map" the arena. Although there are extensive empirical results from this task, there are few models of how it might be accomplished. Moreover, some results from reorientation studies have been different from those of other spatial learning studies. For example, geometric information is typically encoded despite the presence of more informative featural information and this has led to the suggestion that organisms may process shape information in a dedicated "geometric module" (Cheng, 1986; see Cheng & Newcombe, 2005, for a review).

Recently Miller and Shettleworth (2007) proposed an associative model that appears to deal with some of the unusual and interesting results obtained in the reorientation task. They argued that the reorientation task should primarily be viewed as one of operant learning, and that this type of learning is responsible for many of the task's empirical regularities (Miller & Shettleworth, 2007). Miller and Shettleworth proposed an operant version of Rescorla and Wagner's (1972) learning model, and used this model to replicate a large number of experimental results. However, although Miller and Shettleworth's insight about the operant nature of the reorientation task is extremely valuable, there is a serious mathematical problem with their model. As is detailed in the following paragraphs, their model does not implement operant learning in the manner that they intended.

The Miller and Shettleworth Model

In Miller and Shettleworth's (2007) model, geometry and features serve as cues that indicate whether a particular location is a likely source of reinforcement. That is, there is no real distinction between geometric and featural cues apart from there being different sources of information relevant to the same task. As a result, Miller and Shettleworth describe locations (e.g., the four corners of a rectangular arena) in terms of whatever cues are present. Learning in the reorientation task occurs when the associative strengths of the available cues change as a function of a participant's exposure to them (as well as to a participant's exposure to reinforcement). More important, they employ a variant of the Rescorla–Wagner model that views the reorientation task in the context of operant conditioning. As a participant is reinforced for exploring a location that has given cues, it will be more likely to

Michael R.W. Dawson, Marcia L. Spetch, and Brian Dupuis, Department of Psychology, University of Alberta; Debbie M. Kelly, Department of Psychology, University of Saskatchewan.

Correspondence concerning this article should be addressed to Michael R.W. Dawson, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada T6G 2P9. E-mail: mdawson@ualberta.ca

explore that and other locations having those cues and less likely to explore locations lacking them. They modify the Rescorla–Wagner model to reflect the changing probabilities of visiting the different locations. This is done by modifying associative strengths using the Rescorla–Wagner equation, but making this modification contingent on the probability that a particular location is visited, which of course will vary as cues at that location change in associative strength.

To be more precise, let V_L represent the total associative strengths of all of the cues available at some location L at some point in training, and let $\Sigma\, V_L$ represent the sum of all of the V_Ls for all of the locations at the same time. Miller and Shettleworth (2007, their Equation 2) define the probability (P_L) of visiting a particular location L at this point in training as:

$$P_L = \frac{V_L}{\sum V_L}. \tag{1}$$

Miller and Shettleworth (their Equation 3) then use this definition of the probability of visiting a location to formulate an operant version of the Rescorla–Wagner model as:

$$\Delta V_E = \alpha(\lambda - V_L)P_L. \tag{2}$$

In this equation ΔV_E is the change in associative strength of some element E that is available at location L, α is a constant (set to a value of 0.15 in Miller and Shettleworth's simulations), and λ is equal to 1 if the participant is reinforced at L, and to 0 otherwise.

A Flaw in the Miller and Shettleworth Model

In mathematical psychology, there is a long history associated with the type of equation that Miller and Shettleworth (2007) used to define the probability of visiting a location (Gulliksen, 1953; Herrnstein, 1970; Luce, 1959, 1961, 1977; Thurstone, 1930). Thurstone (p. 470) defined the probability of an act leading to successful consequences (P) as $P = s/(s + e)$, where s can be viewed as the strength of a successful response and e can be viewed as the strength of competing responses. Luce (1959, Theorem 3) explored the possibility of creating a numerical scale, which he called the "v-scale", that could be used to represent choice behavior when a number of different choices were possible, as is the case in the reorientation task. Luce's equation is strikingly similar to the probability equation used by Miller and Shettleworth:

$$P_T(x) = \frac{v(x)}{\sum\limits_{y \in T} v(y)}. \tag{3}$$

In Luce's equation $P_T(x)$ is the probability of choosing x from the set T, $v(x)$ is the current strength of x, and the denominator is the sum of the strengths of all of the possible choices from the set. Herrnstein (1970) used similar equations to model the law of effect; that is, to account for the relationship between rates of making particular responses and the various reinforcements that these responses receive.

More important, one requirement in all of the equations cited in the previous paragraph is that all of the elements in the numerator and in the denominator must have positive values. Luce (1959, p. 95) called this the *positiveness condition*. The problem with the Miller and Shettleworth (2007) model is that they violate the positiveness condition when probabilities are calculated in Equation 1 above. This is because their calculation of probabilities is based on associative strength, which can be positive or negative (see e.g., their Figure 1). Because some associative strengths of

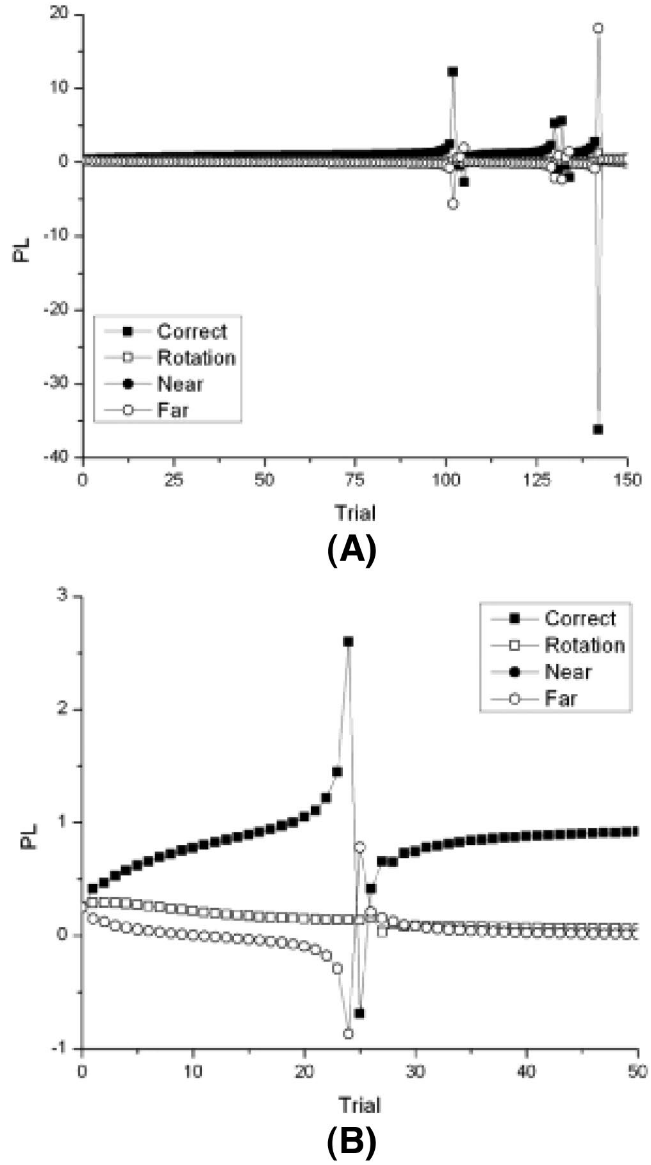

Figure 1. Aberrant behavior of the Miller and Shettleworth (2007) model, using example equations from their paper. In both graphs, "correct" is the probability associated with the correct corner, "near" is the probability associated with the nearest incorrect corner, "rotation" is the probability associated with the incorrect corner diagonally opposite to the correct corner, and "far" is the probability associated with the remaining incorrect corner, as per the terminology used by Miller and Shettleworth. (A) The example they provide, using α of .15 but with training extended to 150 trials. Note several instances where P_L extends beyond the range between 0 and 1, which is impossible if this value represents a probability. (B) Similar aberrant behavior is observed much earlier in the set of equations when α is increased to .60.

elements can be negative, it is possible that the iterative application of their learning rule (Equation 2 above) can cause some values of V_L to become negative. This in turn can produce "impossible" values for P_L, if P_L is to be interpreted as a probability. That is, P_L can become negative or (because the sum of P_Ls over all the locations must equal 1), greater than one. When "impossible probabilities" are used in Equation 2 above, unanticipated changes in associative strengths occur. In short, when strengths of elements change as specified in Equation 2, Equation 1 can violate the positiveness condition.

The aberrant behavior of the Miller and Shettleworth (2007) model can easily be demonstrated. They provided detailed equations for an example reorientation task. We implemented these equations in a Microsoft Excel® spreadsheet. In our first simulation, we replicated their results for the first 20 trials of learning, and then continued training. Small negative probabilities appeared for two of the locations in the reorientation task at Trial 48, and the "probability" of choosing the correct location rose to greater than 1 at Trial 84. Radical shifts in probability were observed at about Trial 100 and then again around Trial 140, as is illustrated in Figure 1A, which demonstrates (for example) that the probability of choosing the correct location drops to -36.22 at Trial 142. In our second simulation, we took their example equations again, but increased α from .15 to .60. This caused the impossible probabilities to be generated much earlier during the simulation (Figure 1B). Clearly Miller and Shettleworth's equations are flawed. Their model performed as well as they reported only because they chose a small enough value for α, and because they examined the results of their equations over only the first 20 learning trials.

Although the mathematics of the Miller and Shettleworth (2007) model are flawed, their operant interpretation of the reorientation task is sound. Fortunately it is possible to use the activation produced by a perceptron to correct the mathematical problem with their model and to implement an operant learning scheme. A perceptron is a particular type of artificial neural network that has a set of input units that are used to encode stimuli, and has one or more output units that are used to represent responses to stimuli (Dawson, 2004, 2005).

As noted above, the mathematical problem faced by the Miller and Shettleworth (2007) model was that violation of the positiveness condition produced impossible probabilities (i.e., $P_L < 0$, or $P_L > 1$). This problem can be solved by making the likelihood of associative strengths being modified a function of a number that is guaranteed to remain between zero and one. One such number is the activity produced by the output unit of a perceptron.

The perceptron is a simple artificial neural network that has a set of input units to represent stimuli, one or more output units to represent responses, and connections (with modifiable weights) from input units to output units (Rosenblatt, 1958, 1962). The output unit can use one of a variety of nonlinear equations to transform the total, weighted signal from the input units into a response (e.g., Dawson, 2004, chapter 10). Perceptrons can be used to model a variety of animal learning tasks by using input units to represent the presence or absence of cues (e.g., Dawson, 2008; Sutton & Barto, 1981).

When the activity of a perceptron's output unit is defined by the logistic equation (which transforms an incoming signal using a sigmoid-shaped function that ranges between 0 and 1), this activity cannot become negative and it cannot exceed 1. Furthermore, this value is based on current associative strengths, as is desired in the Miller and Shettleworth model, because output unit activity depends on the signals sent through existing connection weights. In addition, increases in activity represent increases in expectation of reinforcement. This is because the desired value of output unit activity is analogous to λ in the Rescorla–Wagner model (Dawson, 2008; Gluck & Bower, 1988; Sutton & Barto, 1981).

We have conducted simulations of the reorientation task and have shown that output unit activity can be viewed as an approximation of the likelihood of reward (Dawson, Kelly & Spetch, 2007). When the target location is uniquely identifiable, output activity to it is in the order of 0.90 and activity to other locations is near zero. When ambiguity exists between two geometrically equivalent locations, output unit activity to the correct location and to its geometrically equivalent counterpart is in the order of 0.45. In a square arena, output activity to any of the corner locations is in the order of 0.24. In sum, output unit activation can serve as a replacement for P_L when this value is used to guide operant learning. Simulations of operant behavior have revealed the value of this approach (Dawson, Dupuis, Spetch, & Kelly, 2008). We are currently developing the perceptron model to show how it can be used to simulate other variations of the reorientation task, to test hypotheses about how spatial information may be encoded, and to make novel predictions for tasks that have yet to be empirically explored.

References

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition, 23,* 149–178.

Cheng, K., & Newcombe, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin & Review, 12,* 1–23.

Dawson, M. R. W. (2004). *Minds and machines: Connectionism and psychological modeling.* Malden, MA: Blackwell.

Dawson, M. R. W. (2005). *Connectionism: A hands-on approach.* Oxford, England: Blackwell.

Dawson, M. R. W. (in press). *Connectionism and classical conditioning. Comparative Cognition and Behavior Reviews.*

Dawson, M. R. W., Dupuis, B., Spetch, M. L., & Kelly, D. M. (2008). Simple artificial neural networks that match, exploit, and explore when confronting a multiarmed bandit. *Neural Information Processing: Letters and Reviews,* under editorial review.

Dawson, M. R. W., Kelly, D. M., & Spetch, M. L. (2007, June). Using artificial neural networks to simulate the reorientation task. Poster presented at the 17th annual meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Victoria, British Columbia.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning—An adaptive network model. *Journal of Experimental Psychology—General, 117,* 227–247.

Gulliksen, H. (1953). A generalization of Thurstone's learning function. *Psychometrika, 18,* 297–307.

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

Luce, R. D. (1959). *Individual choice behavior.* New York,: Wiley.

Luce, R. D. (1961). A choice theory analysis of similarity judgments. *Psychometrika, 26,* 151–163.

Luce, R. D. (1977). The choice axiom after 20 years. *Journal of Mathematical Psychology, 15,* 215–233.

Miller, N. Y., & Shettleworth, S. J. (2007). Learning about environmental geometry: An associative model. *Journal of Experimental Psychology: Animal Behavior Processes, 33,* 191–212.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review, 65,* 386–408.

Rosenblatt, F. (1962). *Principles of Neurodynamics.* Washington, DC: Spartan Books.

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

Thurstone, L. L. (1930). The learning function. *Journal of General Psychology, 3,* 469–493.

Received September 21, 2007
Revision received November 1, 2007
Accepted February 12, 2008 ■