Bayesian Models of Instrumental Learning: A Response to Dupuis and Dawson

Noam Miller
Princeton University

Sara J. Shettleworth
University of Toronto

Miller and Shettleworth (2007) used an associative model of instrumental choice to explain a confusing pattern of results in the geometry learning literature. Dupuis and Dawson (in press) identified a structural flaw in the Miller-Shettleworth (MS) model and suggested replacing it with an operant perceptron model which can correctly reproduce some experimental results that the MS model does not. Here we demonstrate that the error in the MS model can be easily corrected without altering any of the model’s predictions by making it stochastic rather than deterministic. In addition, we show that the raw outputs of the perceptron model cannot be interpreted as discriminative choices in an instrumental task without first being normalized. We show that this additional step renders the results of the perceptron model identical to those of the MS model in exactly those cases in which it has been claimed to correctly predict results that the latter cannot.

Keywords: Miller-Shettleworth model, geometry learning, perceptron, associative learning

Recently, we (Miller, 2009; Miller & Shettleworth, 2007, 2008) proposed an associative model of instrumental learning to explain a series of experimental results on geometry learning (see Cheng & Newcombe, 2005, and below). Our model (which we refer to as MS) has been tested against a wide range of data and correctly predicts the outcomes of many, but not all, geometry learning experiments (Cheng, 2008; Cheng, Huttenlocher, & Newcombe, in press). In response to a comment by Dawson, Kelly, Spetch, and Dupuis (2008) we altered the model slightly (Miller & Shettleworth, 2008) without affecting any of its predictions. However, Dupuis and Dawson (in press) have identified a more fundamental flaw in the structure of the MS model that causes it to give nonsensical results under certain parameter regimes. Here, following the suggestions of Dupuis and Dawson, we correct this problem and demonstrate that the revised model gives qualitatively identical results to the original in all cases. In addition, we show that the alternative model suggested by Dupuis and Dawson, based on a type of neural network called a perceptron, also gives qualitatively similar results (to our revised model) in all the situations under which it has been tested.

The MS Model

In a geometry learning (sometimes called “reorientation”) task, the subject is placed into an arena with a distinctive shape, usually a rectangle. A reward is hidden in one corner of the arena and various cues (referred to as features) may be placed at the corners to disambiguate them. The subject must make a choice of which corner to search at and receives the reward if it is correct. In a rectangular arena, corners may either have a long or short wall to their left—and vice versa on their right—and the arena’s geometry therefore differentiates the correct corner and the corner diagonally opposite it (the rotational corner) from the other two (geometrically incorrect) corners.

Studies on a wide range of species (Cheng & Newcombe, 2005) have shown that even when a feature indicates the correct corner, individuals continue to make errors to the rotational corner, implying that they also learn to use the geometry of the arena to locate the reward (e.g., Cheng, 1986; Graham, Good, McGregor & Pearce, 2006; Pearce, Graham, Good, Jones, & McGregor, 2006; Wall, Botly, Black, & Shettleworth, 2004). Surprisingly, several studies have shown that features neither overshadow nor block learning about the geometry (Graham et al., 2006; Hayward, Good, & Pearce, 2004; Kelly & Spetch, 2004a, 2004b; McGregor, Horne, Esber, & Pearce, 2009; Pearce, Ward-Robinson, Good, Fussell, & Aydin, 2001; Wall et al., 2004), which has led to suggestions that geometry learning takes place in a dedicated geometric module (Cheng, 2008). However, in other studies, features can block learning based on the shape of the arena (Gray, Bloomfield, Ferrey, Spetch, & Sturdy, 2005; Horne & Pearce, 2009; Kosaki,
choosing a location \( L \) in which it occurred, \( PL \):

we scaled the change in associative strength of each cue by the probability of animals” (Miller & Shettleworth, 2007, p. 194), we scaled the total change in the associative strength of a cue thus

in which \( \lambda \) represents the sum of the associative strengths only of those cues present at location \( L \) (corresponding to one of the corners of the arena in a geometry learning experiment; note, however, that the model could also be applied to the different options in any operant paradigm). The learning rate parameter \( \beta \) is set to 1, for simplicity (Miller & Shettleworth, 2007).

In attempting to create a deterministic version of the model, which would represent “the mean performance of a large group of animals” (Miller & Shettleworth, 2007, p. 194), we scaled the change in associative strength of each cue by the probability of choosing a location \( L \) in which it occurred, \( P_L \):

\[
\Delta V_A = \alpha(\lambda - V_L), \quad (1)
\]

The total change in the associative strength of a cue thus depended on the summed probabilities of choosing each location containing that cue. The probability of choosing any location \( L \) was given by:

\[
P_L = \frac{V_L}{\Sigma V_L}, \quad (3)
\]

in which \( \Sigma V_L \) represents the summed associative strengths of all locations. Note that if certain cues occur at more than one location (e.g., the cue representing the rewarded geometry in a rectangle) their associative strengths will appear more than once in the denominator of this function.

As Dupuis and Dawson (in press, Appendix) show, multiplying Equation 1 by a location’s choice probability \( (P_L) \) causes the model, under certain parameter values, to give wildly fluctuating associative strengths and choice probabilities outside the range 0 to 1. In an attempt to remedy this problem, we (Miller & Shettleworth, 2008) altered the model’s choice function (Eq. 3) using what Dupuis and Dawson refer to as the “positiveness correction.” As these authors rightly point out, because the structure of Equation 2 remained unchanged, this did not solve the problem. Dupuis and Dawson further suggest that, due to this flaw, the MS model should be abandoned. We, however, believe that the error in Equation 2 can be remedied quite simply and demonstrate below that the corrected model retains all the explanatory power and structural simplicity of the original.

A Revised Model

The error in Equation 2, as Dupuis and Dawson (in press) correctly point out, results from scaling the associative strengths by the choice probabilities \( (P_L) \), which are themselves a function of the associative strengths at each location. The simplest solution to this is to remove \( P_L \) from the equation, reverting it to the form of Equation 1. However, we still require that the associative strengths of cues only change when they are experienced. Thus, the choice probabilities (Eq. 3) will still determine which location is chosen, which in turn will determine whether or not a given cue’s associative strength is updated. This is now done stochastically.

The revised model then operates as follows.

At the start of each trial, the choice probabilities for each location are calculated, using Equation 3. To keep the choice probabilities within the required range, we retain a form of “positiveness correction”: if \( P_L \) of any location is greater than 1 or smaller than 0, it is set to 1 or 0, respectively. Note, however, that the positiveness correction only ensures that the probability of choosing any given location will fall between 0 and 1. The sum of the probabilities for all locations can still exceed 1. In most geometry learning studies, individuals make only one choice per trial (indeed, in most operant tasks a trial is defined by the occurrence of a single choice), and we therefore normalize the choice probabilities of all locations such that they sum to 1 before determining the simulation’s choice on any trial:

\[
P_L = \frac{V_L}{\Sigma V_L}, \quad (4)
\]

in which \( V_L \) represents the summed associative strength of the cues at any location \( L \). This is comparable with asserting that the subject considers all locations before making a choice. Though this may not always be true, we consider it the simplest assumption and likely the optimal strategy. The choice made by the simulation on that trial is determined randomly, weighted by the relative choice probabilities of all the locations, such that the distribution of choices reflects the relative attractiveness of each location. On making a choice, the associative strengths of all the cues present at that location are updated using Equation 1. The updated associative strengths are subsequently used to determine the choice on the next trial, using Equation 4, and so on. Because the choice stage of this process is stochastic, we repeat the simulation many times and present averaged results.

1 Computationally, for each trial, a random number in the interval [0, 1) is generated and compared with the discrete cumulative distribution of \( P_L \) across all locations, \( f(P_L) \). For example, for a simulation with four equally attractive locations, \( f(P_L) = [0.25, 0.5, 0.75, 1] \). The simulation then chooses the location corresponding to the smallest member of \( f(P_L) \) larger than the random number (e.g., if the random number was 0.6, the simulation would choose the third location in the example above).
The revised model accurately reflects the intentions (and verbal description) of the original MS model but does not suffer from any of the mathematical flaws identified by Dawson et al. (2008) or Dupuis and Dawson (in press). Below we simulate the two scenarios also simulated by Dupuis and Dawson and show that the revised model makes the same predictions, qualitatively, as the original MS model.

**The Perceptron Model**

After explaining the errors in the MS model, Dupuis and Dawson (in press) suggest an alternative model to simulate geometry learning experiments. First, they suggest using a logistic choice function, which cannot return probabilities outside the range 0 to 1. In the notation of the MS model this becomes:

\[ P_L = \frac{1}{1 + e^{-V_L}}. \]  

(5)

This suggestion can easily be incorporated into the revised MS model, by replacing Equation 4 with Equation 5, rendering the structure of the model identical to that of the perceptron model (Dawson, 2008). However, the values returned by this choice function, like those given by Equation 4, must still be normalized to obtain actual choice probabilities. In other words, though \( P_L \) for each location is now guaranteed to be between 0 and 1, the sum of the probabilities over all locations may still exceed 1. If subjects consider all alternatives before making a single choice, the final probability of choosing any location L should be \( \frac{p_L}{\sum_i p_i} \). Below we run each of our simulations with both choice rules and show that they give qualitatively similar results.

In addition to the logistic choice rule, Dupuis and Dawson (in press) suggest using a perceptron, a simple neural network, to model the reorientation task. They suggest that the perceptron, which has been shown to correctly reproduce some geometry learning data (Dawson, Kelly, Spetch, & Dupuis, 2010), can correctly predict Horne and Pearce’s (2010) experimental results, which the MS model fails to do. Below we show that the perceptron model does not correctly reproduce these results if its outputs are interpreted in a manner consistent with the experimental paradigm and that the perceptron and revised (and original) MS models both predict the same outcome, incorrectly.

**Simulation Results**

The simulations below were performed similarly to those reported in Miller and Shettleworth (2007, 2008), with the exception that the revised version of the model, presented above, was used. Simulations were run using both Equations 4 and 5 as the choice rule. Each simulation was trained for 30 time steps and the results of 500 replications of each simulation were averaged together. In all simulations the saliences of all cues, \( \alpha \), were set to 0.15 (as in Miller & Shettleworth, 2007). However, as the aberrant behavior of the original MS model only emerged at high values of \( \alpha \), we also tested each simulation at a range of values of \( \alpha \) up to 1, to ensure it did not give nonsensical results.

Wall et al. (2004, Experiment 3) attempted to block learning about the geometry of their rectangular arena with a large black feature placed at the correct (rewarded) corner. Individuals in the Blocking group were first trained to associate the feature with a reward in a square arena (in which all corners provide identical geometric information). Then, both the Blocking and a Control group were trained to find the reward in one corner of a rectangular arena containing the same feature. At the end of training, both groups were tested in the rectangular arena without the feature. If previous experience with the feature blocked learning about the geometry in the Blocking group, they would be expected to do worse at test than the Control group. However, the authors found no difference between the groups.

This result is correctly predicted by the original MS model (Miller & Shettleworth, 2007), though the model misbehaves if \( \alpha \) is set higher than about 0.7 (Dawson et al., 2008; Dupuis & Dawson, in press), as well as by the perceptron model (Dupuis & Dawson, in press). We simulate the same experiment using the revised MS model and show that it also reproduces the results.

The simulation involves four cues: a context cue present at all locations (B); the geometry of the correct and rotational corners (G); the geometry of the incorrect corners (W); and the feature at the correct corner (F). The associative strength of the feature (\( V_F \)) in the Blocking group is initialized at 0.3, to represent their initial phase of training; the associative strength of the context cue (\( V_B \)) is started at 0.1; all other cues have initial associative strengths of 0 (see Miller & Shettleworth, 2007, for details of the simulation procedure).

Figure 1 shows the results of the simulation using the original MS choice rule (Eq. 4). Both the associative strengths (Figure 1 A, C) and the choice probabilities (Figure 1 B, D) are virtually identical to those reported by Miller & Shettleworth (2007, Figure 1) and fit the experimental data well (Wall et al., 2004, Figure 4). When the model is tested (by removing the feature), the Blocking group is predicted to choose the geometrically correct corners 87% of the time; the Control Group 89% of the time. These results are also congruent with those of the original model and the experimental data. The simulation also correctly predicts the lack of blocking when run using the logistic choice rule (Eq. 5), though the results for both groups are far less extreme (Blocking Group 52.5% correct; Control Group 53.4% correct).

The revised model is also robust to changes in the value of \( \alpha \). For example, when run with \( \alpha = 0.65 \) (for all cues) for 10 training trials, the simulation predicts perfect performance (100% correct) by both groups. Interestingly, if \( \alpha \) is increased further or the simulation is run for longer, the performance of both groups at test declines toward chance levels (additionally, all the associative strengths begin to fluctuate when \( \alpha \) exceeds about 0.85 but this does not affect the choice probabilities). This is because, with increased or faster learning, the feature—which is the best predictor of reward during training—usurps most of the associative strength, causing the associative strength of the correct geometry to decrease. At test, with the feature removed, the model suggests that the weakening of geometric cues will lead to worse performance by both groups than if learning was slower or more brief (formally, all the cues other than the feature become conditioned inhibitors and, at test, Eq. 4 returns a probability of 0 for all
locations, implying random choice). In other words, a high $\alpha$ or a large number of training trials will eventually result in the feature blocking (or overshadowing, in the Control group) geometric cues. A similar prediction was made by Miller and Shettleworth (2007) and experimentally confirmed by Horne and Pearce (2009; see also Kosaki et al., in press and McGregor et al., 2009).

Horne and Pearce (2010), Experiment 2a

Next we simulate the superconditioning experiment conducted by Horne and Pearce (2010, Experiment 2a), also modeled by Dupuis and Dawson (in press). In a superconditioning paradigm, a cue gains additional associative strength by being paired with a conditioned inhibitor. In the first phase of the experiment, Horne and Pearce trained two groups of rats to find the escape platform in a rectangular water maze using only the geometry of the maze (platforms were located at both geometrically correct corners). In Phase 2, identical trials to Phase 1 were interspersed with trials on which a feature was present at both geometrically correct corners and there was no platform in any corner. This training regime should result in the feature becoming a conditioned inhibitor. In the third and final phase, rats in the Experimental group were trained in the rectangular arena with the feature and a platform present at both correct corners. In this phase, previously learned geometric cues, paired with the (inhibitory) feature, are expected to gain additional associative strength. A Control group was trained with a novel feature at the correct corners. When tested in the absence of the features at the end of training, rats from the Experimental group spent more time in the geometrically correct corners than rats from the Control group, demonstrating superconditioning of the geometric cues by the inhibitory feature.

Horne and Pearce (2010) attempted to simulate their results using the original MS model and found that the model incorrectly predicted a lack of superconditioning, as does the revised model presented here. Figure 2 shows the results of the revised model for the three phases of the experiment. The results are qualitatively identical to those presented by Horne and Pearce (2010, Figure 11) and similar to the simulation of this experiment by Dupuis and Dawson (in press, Figure 5). In the final phase, the model predicts that the correct geometry will gain additional associative strength in the Experimental, but not the Control, group. As a result, at test, Equation 4 returns higher raw probabilities for the correct corners in the Experimental group (Experimental 1.0; Control 0.94). However, when these values are normalized to give actual choice probabilities the model predicts a slight advantage for the Control group (Experimental 0.90 correct; Control 0.95). These results are comparable with those obtained by Horne and Pearce (2010) using the original MS model. Horne and Pearce’s experiments were performed in a water maze, in which subjects make multiple choices until they locate the submerged platform (the reward), and the authors therefore used the multiple-choice version of the original MS model (see Miller & Shettleworth, 2007) to simulate their data. We note that using a stochastic version of the multiple-choice model, constructed along the same lines as the revised single-choice model presented above, gives qualitatively identical results to those presented here.

Dupuis and Dawson (in press) claim that the perceptron model, unlike both versions of the MS model, is able to correctly reproduce...
Horne and Pearce’s (2010) experimental results. The authors present the activations—the outputs of Equation 5—of their perceptron model at testing: 0.998 for the Experimental group and 0.970 for the Control group. However, as noted above, the outputs of Equation 5 (or of Eq. 4) cannot be considered choice probabilities in a geometry learning task, because the sum of the responses to the different locations may exceed 1 (as in the present case). Unfortunately, the authors do not report the activations to the incorrect corners. We therefore recreated the perceptron model to explore whether, when activation levels are correctly converted into choice probabilities, it still predicts superconditioning in this scenario. Using the same parameter values reported by Dupuis and Dawson, we find the following activation levels at test: Experimental group, 0.998 to the geometrically correct corners, 0.090 to the incorrect corners; Control group, 0.985 to the geometrically correct corners, 0.068 to the incorrect corners. Note that the activations to the correct corners are higher in the Experimental group, as reported by Dupuis and Dawson. However, the activations to the incorrect corners are also higher and, when these values are normalized to extract choice probabilities, the simulation predicts superconditioning in this scenario. Using the same parameter values reported by Dupuis and Dawson, we find the following activation levels at test: Experimental group, 0.998 to the geometrically correct corners, 0.090 to the incorrect corners; Control group, 0.985 to the geometrically correct corners, 0.068 to the incorrect corners. Note that the activations to the correct corners are higher in the Experimental group, as reported by Dupuis and Dawson. However, the activations to the incorrect corners are also higher and, when these values are normalized to extract choice probabilities, the simulation predicts 92% correct choices in the Experimental group versus 94% correct in the Control group. In other words, the perceptron model does not reproduce the superconditioning effect observed by Horne and Pearce (2010). Indeed, the higher activations but lower choice probabilities in the Experimental group are identical to the results of the revised (and original) MS model.

**Activations and Choice Probabilities**

The discrepancy between the activations of the perceptron (the outputs of Eq. 5) and normalized choice probabilities reveals the problem with using the perceptron to simulate geometry learning. Perceptrons return conditional probabilities, that is, likelihoods of response given a certain set of input stimuli. However, they do not explicitly compare options (locations) with each other. Perceptron activations have been considered comparable with actual response rates (Dawson, Dupuis, Spetch, & Kelly, 2009) but, even then, the model is presented with a single option at a time and returns a probability of response to that option only. Other options are evaluated independently (Dawson, 2008).

To clarify why this is important when simulating geometry learning results, we perform the following thought experiment: A subject is trained to find a reward in two geometrically equivalent corners of a rectangular arena, both marked by identical features (some geometry learning experiments have used a similar training setup; e.g., Hayward et al., 2004). In this situation, geometric cues and the features are equally good predictors of the reward location and, according to the MS model, should both gain associative strength. After training, the subject is tested in the same rectangular arena either with no feature present or with one feature placed at one of the geometrically incorrect corners.

We simulate this scenario using both the revised MS model and the perceptron, using the same parameter values as for the simulations above. The MS model predicts that when tested in a bare rectangular arena subjects should search in the geometrically correct corners 84% of the time. When the feature is added to one of the geometrically incorrect corners the model predicts that subjects search at that corner 25% of the time and at the geometrically correct corners only 68% of the time. In other words, the attractive feature reduces the probability of geometrically correct responses. Empirical results supporting this prediction have been reported in situations in which features are moved between training and testing (e.g., Cheng, 1986; Kelly, Spetch, & Heth, 1998; Kelly & Spetch, 2004b; Sovrano, Bisazza, & Vallortigara, 2003).
We next simulate the same scenario using the perceptron model. When tested in a bare rectangular arena, the perceptron returns activations of 0.81 to the geometrically correct corners. However, as explained above, this value depends only on the associative strengths of the cues at those locations. When the attractive feature is added to one of the geometrically incorrect corners, the perceptron still returns an activation of 0.81 to the geometrically correct corners. In other words, interpreting the activations of the perceptron as choice probabilities—as Dupuis and Dawson (in press) do—suggests that adding the feature to a geometrically incorrect corner will not change the likelihood of visits to the geometrically correct corners. As noted above, this is contradicted by experimental results. Of course, the activations of the perceptron to the geometrically incorrect corners do change (0.03 when the arena is bare; 0.21 when the feature is present), but these changes cannot affect the activation to the correct corners. If, as we suggest above, the activations are normalized to obtain choice probabilities, the perceptron’s predictions are identical to those of the revised MS model (bare arena: 96% geometrically correct; with feature: 87% geometrically correct, 11% to the corner with the feature).

**Conclusion**

Dupuis and Dawson (in press) identified a flaw in the mathematical formulation of the original MS model (Miller & Shettleworth, 2007, 2008). We show above that this error in the model can be easily corrected by making the model stochastic rather than deterministic and retaining both its performance function (Eq. 4) and its learning rule (Eq. 1). The revised model gives qualitatively identical results to the original MS model in all the cases in which we have tested it and is as straightforward to interpret as the original (code for the revised model is available from the corresponding author). In addition, we show that the perceptron model cannot be used in its current form, as suggested by Dupuis and Dawson (in press), to simulate geometry learning experiments, because the outputs of the model do not correspond to the kinds of discriminative choices that subjects in these experiments are required to make. It is possible to convert the outputs of the perceptron into choice probabilities but, when this is done, the predictions of the model align perfectly with those of the MS model (see also Dawson, 2008).

The MS model has been empirically tested several times (e.g., Horne & Pearce, 2009, 2010, 2011; McGregor et al., 2009; Miller, 2009; Sturz & Kelly, 2009) and the model fails to reproduce some of these data, including the superconditioning task simulated above (Horne & Pearce, 2010). Additional work is required to determine whether the perceptron model can correctly simulate the results of some of the studies in which the MS model fails. This work must begin, in our opinion, by showing that the outputs of the perceptron can be converted into discrete choices of the kind that animals in geometry learning tasks perform. Such a demonstration would indeed suggest that the perceptron provides a more useful framework for understanding these data than the MS model.

---

2 Perceptron simulations were run using Michael Dawson’s “Rosenblatt” program, available at [http://www.bcp.psych.ualberta.ca/~mike/Software/Rosenblatt/index.html](http://www.bcp.psych.ualberta.ca/~mike/Software/Rosenblatt/index.html). Simulations were run using the same parameters as Dawson and Dawson (in press) used for their superconditioning simulation.

---

**References**


of Comparative Psychology, 118, 384–395. doi:10.1037/0735-7036.118.4.384


New Editors Appointed, 2015–2020

The Publications and Communications Board of the American Psychological Association announces the appointment of 6 new editors for 6-year terms beginning in 2015. As of January 1, 2014, manuscripts should be directed as follows:

- Behavioral Neuroscience (http://www.apa.org/pubs/journals/bne/), Rebecca Burwell, PhD, Brown University
- Journal of Applied Psychology (http://www.apa.org/pubs/journals/apl/), Gilad Chen, PhD, University of Maryland
- JPSP: Interpersonal Relations and Group Processes (http://www.apa.org/pubs/journals/jsp/), Kerry Kawakami, PhD, York University, Toronto, Ontario, Canada
- Psychological Bulletin (http://www.apa.org/pubs/journals/bul/), Dolores Albarracin, PhD, University of Pennsylvania
- Psychology of Addictive Behaviors (http://www.apa.org/pubs/journals/adb/), Nancy M. Petry, PhD, University of Connecticut School of Medicine

Electronic manuscript submission: As of January 1, 2014, manuscripts should be submitted electronically to the new editors via the journal’s Manuscript Submission Portal (see the website listed above with each journal title).

Current editors Mark Blumberg, PhD, Steve Kozlowski, PhD, Arthur Graesser, PhD, Jeffry Simpson, PhD, Stephen Hinshaw, PhD, and Stephen Maisto, PhD, will receive and consider new manuscripts through December 31, 2013.