“Drink way too much coffee and buy a desk that doesn't collapse when you beat your head against it.”

--Douglas Adams, on getting ideas

University of Alberta

The Cognitive Science of Reorientation

by

Brian Dupuis

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Psychology

Fall 2012
Edmonton, Alberta

Dedicated to the memory of my friend and colleague, Matthew Ian Helgesen (1983-2011).

Abstract

This work stands as an example of “synthetic methodology” in psychological research. Synthetic methodology involves building a model, seeing what it can and cannot do when placed in interesting environments, comparing this behaviour to real-world subjects for parallels and discrepancies, and then examining the model for insight and theoretical advancement. This methodology is employed here in the context of a common spatial-learning “reorientation task”. Motivated by the discovery of critical flaws in a popular model for this reorientation task, we develop a synthetic neural network model as an alternative, and explore its behaviour in novel tasks, as well as the mathematical consequences of adopting such a formalism. These behaviours lead us to question assumptions underlying normal reorientation research. We devise a new method of collecting human data in spatial tasks, and use this method to compare the neural network to human subjects, in the style of comparative cognition.

Acknowledgement

I would like to thank Marcia Spetch and Patricia Boechler for their aid throughout this project, as well as Nolan Thomas Denman, Gerry Leenders, Michel Fiallo-Perez and Vadim Kasim for helpful discussion during preparation of Chapter 1. I would like to extend special thanks to Erik deJong for his invaluable help in the creation of the software described in Chapter 3; with his help, Lucio Gutiérrez, Danielle Baron (née Lubyk), and I learned much about each other's respective disciplines, and gained a very useful tool out of it. Finally, I would like to offer an enthusiastic and personal thank you to my supervisor, Michael R.W. Dawson, for supporting me in many ways through what has been – by far – the most difficult and trying time of my life. Thank you all.

Table of Contents

List of Tables

List of Figures

Introduction

This thesis is representative of the collection of work I have accomplished in two years while training as an apprentice cognitive scientist. Each of its chapters is a version of a full academic paper at various stages of the publication process; while each stands on its own, taken as a whole, the chapters reveal a broader pattern. Together they reflect an approach for conducting cognitive science that can be very productive and informative, but is not, I feel, common practice among the research community at large.

This specific approach has four of components, each relying on a different, but related, set of skills. The first component is *formal*, where a new model is created based on a mathematical understanding of existing theory. The second involves a firm understanding of what each new formalism means, *theoretically* – that is, a design decision made during formalization will be founded on several tacit assumptions, of which its author should not be unaware. In addition to the typical understanding of current theory used to guide the formation of new models, this second component emphasizes the implicit theory that emerges from the model itself. The third component is the ability to let go of cherished designs, and to *improvise* new ones out of what is at hand – a sort of methodological *bricolage*, if you will. Finally, the fourth component is to recognize that the models one makes are, themselves, some of the richest sources of inspiration one can have for developing and refining both novel theoretical claims, but also for designing and testing new experiments. Rather than only analyzing experimental results to develop ideas, one can instead *synthesize* a new, plausible model first, and use its behaviour to inform future research. If one is open to recognizing when models make surprising predictions, and design new experimental methods if none currently exist to test these claims, one may find unexpectedly fruitful research, even from a null result.

The purpose of this thesis is to illustrate this approach to conducting cognitive science in the context of a particular research domain, the study of navigation and reorientation – the process by which an agent finds its way after becoming lost. The thesis uses this as a framework to provide examples of each of the methodological components in action, as well as how these components can be cooperatively synthesized to provide insights into this research domain. It will progress as follows:

The first chapter, *The Miller-Shettleworth Model's Empirical Difficulties*, begins by identifying a serious mathematical shortcoming in an influential model of associative learning, developed in the context of reorientation. We trace the source of this flaw within two different experimental paradigms, and eventually identify that the problem lies in the form of the model itself – it cannot be saved in its current form. We then propose an alternative model – a revised artificial neural network model with a novel, psychologically-plausible learning algorithm, dubbed an *operant perceptron* – expressly to avoid these formal shortcomings with the Miller-Shettleworth model. This work provides a solid example of the formal component of practical cognitive science – and also shows that when this component is not given appropriate attention, the result is a fundamentally flawed model.

With an alternative to the Miller-Shettleworth model in hand, we must now explore the consequences of adopting this new model's formalism, in accordance with the theoretical component. This begins by going beyond model evaluation as typically practiced in experimental psychology, and instead looks to link the model to well-established mathematics. This is the point of Chapter 2, *The Equilibria of Perceptrons for Simple Contingency Problems*. A critical property of the operant perceptron model is that, in essence, it learns to estimate reinforcement contingencies. Due to the operant perceptron's non-zero response probabilities to non-reinforced locations, the perceptron convergence theorem applies, just as it does to the standard perceptron: both models will, over time, achieve identical equilibria on any problem they can solve. How do these equilibria compare to the established mathematics of this field, found in contingency theory? This chapter elaborates on this formal link and proves the connection.

Chapter 3, *fAARS-Lite: An Open Platform For Investigating Spatial Tasks In Humans* reflects another consequence of the model's formalism. As we thought about the consequences of the model, we also thought about how to test its novel claims in live agents – and found that existing tools to do so often had barriers to entry, which made investigating non-standard tasks difficult. Rather than adjust to the needs of existing tools, we developed our own – fAARS-Lite -

designed to accommodate a wide range of geometry-learning tasks with relatively little computer expertise required. In addition to being simple to employ, the fAARS-Lite platform was designed to be as generalizable as possible – with logic similar to a Turing machine, assembled out of very basic elements in a graphical environment. This ease and generalization should allow others to take the platform and improvise novel environments of their own. Rather than develop the perfect tool for one job, we created an environment inspired by the improvisational component – and as the chapter discusses, this environment has already proven useful in fields unrelated to its original inception (computer science and comparative cognition).

Chapter 4, *Get Out Of The Corner: The Effect Of Location Type And Number On Perceptron And Human Reorientation* is, in many ways, the unison of the earlier methodological components, and a microcosm of the method as a whole, in the context of a standard psychological experiment. In this chapter, we relay the results of our explorations with the operant perceptron model. These simulations force us to question a few tacit assumptions underlying many reorientation experiments, and the results are surprising. We then employed fAARS-Lite to create an analogous task for human subjects, and conducted an experiment to see if the perceptron's predictions were confirmed or rejected. This not only provides a direct example of the "model-informing-experiment" nature of the synthetic component from above, but also produces results of its own. These particular results, we feel, will be of interest to the reorientation research community as a whole. However, we also believe that the results of Chapter 4 – combined with the case studies provided by Chapters 1, 2 and 3 – illustrate the advantages of this multi-component methodology for conducting cognitive science.

Chapter 1

The Miller-Shettleworth Model's Empirical Difficulties

Brian Dupuis and Michael R.W. Dawson

Department of Psychology, University of Alberta

Author Note:

A version of this chapter has been submitted for publication. Dupuis & Dawson, 2012. Journal of Experimental Psychology: Animal Behaviour Processes.

The Reorientation Task

The ability to navigate around in the world is fundamental to nearly every mobile creature. One aspect of navigation that has been studied extensively is reorientation, defined here as an agent's ability to locate a previously learned position when disoriented. At its simplest, reorientation requires an agent to navigate to a previously-learned location based upon available environmental cues, such as information about the shape of the environment ("geometric" cues) and about landmarks present in the environment ("feature" cues). Such behaviour is typically studied experimentally with what has become known as the "reorientation task" (Cheng, 1986), which is described in more detail below. Such experiments have shown that agents, even those from dramatically different species that range from ants to humans, exhibit certain empirical regularities (reviewed in Cheng & Newcombe, 2005).

In a typical reorientation task (Cheng, 1986), an agent freely explores a rectangular arena, with each corner distinguished from the others through some combination of geometric and featural cues. If the agent approaches the corner the experimenter has deemed "correct" - for instance, the corner containing a unique feature like a colored panel, with a long wall on its left - then the agent is reinforced, otherwise no reward is offered. After repeated trials to learn that this location is correct, the agent is disoriented and placed in a new arena in which the feature cues are placed in conflict with the geometric cues - in the above example, the panel is now in a corner with a long wall on its right. It might be plausibly predicted that, in this new arena, agents will move towards the feature, which was the only unique predictor of reward during training. Curiously, while agents will approach the corner with the unique local feature, agents will also frequently choose locations matching the original geometry - the original corner and its geometric equivalent - even though this geometry was not always reinforced during training, and neither corner currently possesses the reinforced local feature. The agent only occasionally follows the feature instead of the geometry, despite the feature uniquely identifying the correct corner every time during training. The exact proportion of responses that follow the feature as opposed to the geometry varies somewhat depending on the size of the arena and the agent in question (for example, see Cheng & Newcombe, 2005; Chiandetti & Vallortigara, 2008), but the general pattern remains consistent.

Recent attempts to provide a theoretical explanation for reorientation task regularities have taken the forms of associative models of an agent's choice behaviour, such as artificial neural networks (Dawson, Kelly, Spetch, & Dupuis, 2010), or as mathematical models in the tradition of Rescorla and Wagner (1972). Such models attempt to explain reorientation by modelling each possible location as a collection of geometric and feature cues that compete for associative strength, with the model's response to a location's particular pattern of cues reflecting its choice behaviour. For example, in the Dawson et al. (2010) network model, the response generated by the network to a particular corner (i.e. to a particular set of cues) was the network's prediction of the probability of being rewarded if that corner was visited. In such a model, geometric cues affect reorientation behaviour because geometric cues and feature cues are learned independently, and geometric cues alone have sufficiently high associative strength to dictate a response, even though these cues are reinforced only for some locations.

The Miller-Shettleworth Model

One associative model of reorientation (Miller & Shettleworth, 2007, 2008) that is of particular interest is an extension of the classic Rescorla-Wagner model of associative learning (Rescorla & Wagner, 1972):

$$\boldsymbol{\Delta V = \alpha \cdot \beta \cdot (\lambda - \Sigma V)} \tag{1-1}$$

In the Rescorla-Wagner model, the change in associative strength ΔV between conditioned stimulus (CS) and unconditioned stimulus (US) is defined by the difference between the magnitude of the US, represented by λ, and magnitude of the current associative strength, represented by V. This difference is scaled by the CS' inherent salience α and by the learning rate related to the US, β. This model is inherently one of classical conditioning. At every iteration, the equation updates associative strengths for all presented cues simultaneously; the agent's choice to respond (or not) is not part of this formulation of learning (Dawson, 2008).

Miller and Shettleworth (2007) convincingly argue that because the agent only chooses one location at a time, and only receives reinforcement or feedback at a chosen location, reorientation is more properly considered an operant task, with reinforcement contingencies based upon the agent's particular pattern of choices. Thus, they modify Equation 1-1 to include a measure of the agent's probability of choosing a given location:

$$\Delta \boldsymbol{V_E} = \boldsymbol{\alpha_E} \cdot \boldsymbol{\beta_L} \cdot (\boldsymbol{\lambda_L} - \mathbf{V_L}) \cdot \boldsymbol{P_L} \qquad (1\text{-}2)$$

Here, the change in associative strength of each cue (or "element", E) is updated using the sum of the associative strengths of all cues at a given location (that is, $V_L=(\Sigma V_E)_L$), and scaled by a term representing the probability of choosing that location. In its original version (Miller & Shettleworth, 2007), this model defined the probability of choosing a given location P_L as the relative associative strength of the location in question compared to the total associative strength at every possible location:

$$\boldsymbol{P_L} = \frac{\boldsymbol{V_L}}{\boldsymbol{\Sigma V_L}} \qquad (1\text{-}3)$$

However, Dawson, Kelly, Spetch, and Dupuis (2008) identified a serious problem when probability is defined using Equation 1-3. They demonstrated that, using an example reorientation problem taken from Miller and Shettleworth (2007), a model that used Equations 1-2 and 1-3 can produce values of P_L that fall outside the range of 0 to 1, and thus cannot be considered "probabilities". In response to the flaw identified by Dawson et al. (2008), Miller and Shettleworth (2008) revised their model of reorientation. Miller and Shettleworth replaced Equation 1-3 with a new term for the relative net *attractiveness* of a location. In the modified model, Miller and Shettleworth defined "net attractiveness of a location" r_L as the sum of the associative strengths of the cues at that location if that sum is positive, or as 0 if that sum is not positive (that is, $r_L=V_L*H_0(V_L)$, where H_0 is the Heaviside step function with threshold 0). P_L became the relative net attractiveness of each location compared to the total relative net attractiveness at every location:

$$\boldsymbol{P_L} = \frac{\boldsymbol{r_L}}{\boldsymbol{\Sigma r_L}} \qquad (1\text{-}4)$$

They reasoned that by replacing Equation 1-3 with Equation 1-4, P_L will always fall within the acceptable range for probability, and they presumed that the theoretical issues with their model had been resolved.

Below, we show that even after this modification, the model still has several underlying problems that render it inadequate for modelling learning in the reorientation task or related associative tasks. The purpose of the current paper is to demonstrate these problems empirically, and explain their source in the model's equations. This analysis of their revised model will also reveal why the errant behaviour identified by Dawson, Kelly, Spetch and Dupuis (2008) emerged in the original version of Miller and Shettleworth's model. Finally, we propose an alternative model to rectify the situation.

By convention, we will refer to models with the form of Equation 1-2 as the "M-S model" regardless of the equation used to compute P_L. The model's original presentation (Miller & Shettleworth, 2007), consisting of Equations 1-2 and 1-3, is referred to as "M-S 2007," while the revised form (Miller & Shettleworth, 2008) using Equations 1-2 and 1-4 is denoted as "M-S 2008". The correction employed by replacing Equation1-3 with Equation 1-4 is referred to as the "positiveness correction", because unlike V_L, r_L is defined in such a way that it is always positive and never negative (as negative sums are artificially set to 0). Finally, the parent model, Rescorla and Wagner's (1972) Equation 1-1, is the "R-W model".

Demonstration of Problems

Miller and Shettleworth (Miller & Shettleworth, 2008) revised their original model to address the fact that it could generate impossible probabilities (Dawson et al., 2008). In this section, we note how the behaviour of the revised model still exhibits important difficulties. We consider two such examples, the first involving a superconditioning paradigm, and the second involving a standard reorientation task. The purpose of the current section is to describe empirical problems with the Miller and Shettleworth model. A later section will provide a detailed discussion of why these difficulties arise.

Superconditioning

A prediction of the R-W model is that "superconditioning" can occur. Superconditioning exists when excitatory cues produce stronger responses after discrimination training if they are paired with an inhibitory cue, compared to a control condition in which the excitatory cues are paired with a neutral cue during training. The presence of an inhibitory cue during discrimination training increases the difference between λ and ΣV, allowing for a greater change in associative strength.

Horne and Pearce (2010, experiment 2) investigated if one could observe superconditioning in the context of geometric and feature cues in a paradigm that can be viewed as an extension of the reorientation task. In this experiment (summarized in Figure 1-1), rats were trained to associate a particular set of geometric cues with reinforcement, but only when a particular feature was absent. That is, in Stage 2 of Figure 1-1, rats are reinforced when a particular set of cues are present at a location (A+), but are not reinforced when those cues are accompanied by an additional cue (AX-). Following this training, the experimental group of rats received reinforcement in the same location with both sets of cues present (AX+), while a control group received reinforcement in that location when the original cues are paired with a novel, neutral cue set (AY+).

Because of their attempt to model animal data using the M-S 2008 model, Horne and Pearce defined the task as consisting of a "correct geometry" cue *G*, an "incorrect (wrong) geometry" cue *W*, a context cue common to every location *B*, and two feature cues (one inhibitory feature present during discrimination training and with the experimental group, *F*, and one neutral feature only present for the control group, *N*)[1]. A summary of these cues as presented in Horne and Pearce's experiment is found in Figure 1-1.

[1]Horne and Pearce employ a different notation, instead using G_C and G_i to refer to correct and incorrect geometric cues, and *F* to refer to any feature cue (despite two being used in the experiment). For consistent terminology across the two experiments, and to avoid confusion about whether a feature cue had prior associative strength (as it would have, during superconditioning) or if it is a novel, neutral cue (as it would be, during controls), we adopt the *G, W, F* and *N* notation described here.

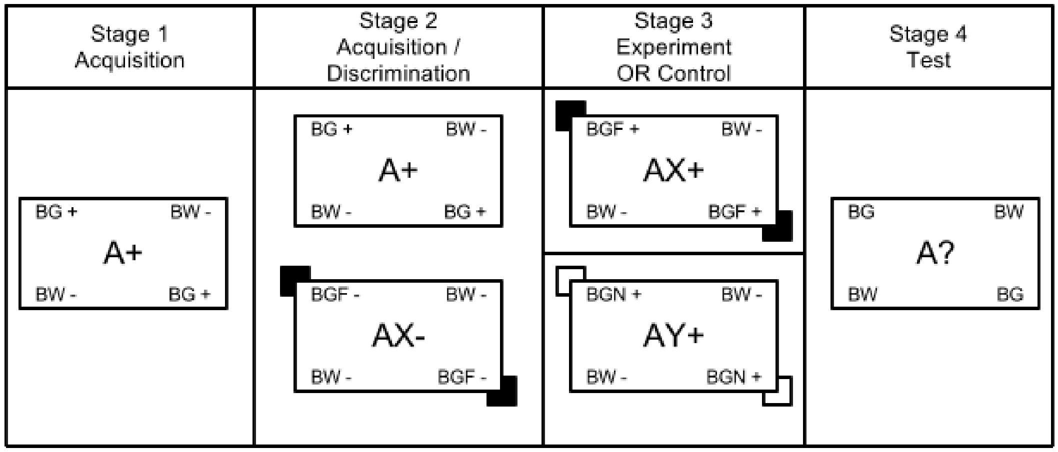

Figure 1-1. A schematic overview of the Horne and Pearce (2010) superconditioning task, adapted from that paper, with stages and cue types labelled. B=”base”, G=”correct geometry”, W=”incorrect geometry”, F=”feature cue”, N=”neutral feature cue”. A + indicates the location is reinforced, a - indicates no reinforcement.

Horne and Pearce’s (2010) results show that the experimental group chose the “correct” corner with greater frequency than did the control group - that is, the rats showed evidence of superconditioning. However, when simulating the same experiment with the M-S 2008 model, Horne and Pearce obtained the opposite result: the probability of choosing the correct corner in the experimental group was 0.92, and the probability of choosing the correct corner in the control group was 0.95. They attributed this failure to display superconditioning to an artificial inflation in the value of the *B* cue, and attempted to solve the problem by manipulating the salience of this cue, but “in all of the simulations that [they] conducted, however, this manipulation does not permit superconditioning to be predicted” (p. 393).

Reorientation

The M-S 2007 model was originally demonstrated using a standard reorientation task (Wall, Botly, Black, & Shettleworth, 2004, Experiment 3, see Figure 1-2 below). Within this task, rats are required to locate food in one corner of a geometric arena - here, the Correct corner, with a particular set of geometric properties *G* and a unique feature *F* (along with a general context cue which Miller and Shettleworth labeled *B*, representing the bowls at each location that the rats searched for food). During this phase of exploration and learning, the rats are also exposed to a different set of geometric properties learned to be wrong, *W*, which never contain reinforcement. In later phases of this experiment, the configuration of cues changes and the rats' behaviour is monitored, but for the purposes of the current paper, we concentrate on this training phase.

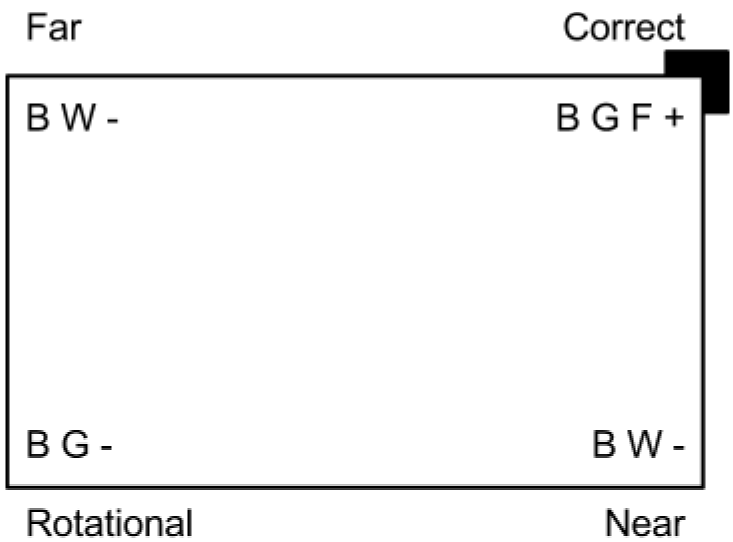

Figure 1-2: Schematic overview of the Wall et al.(2004) reorientation task training phase traditionally used in discussions of the M-S model, with locations and cue types labelled. B="Bowl", G="correct Geometry", F="Feature", W="Wrong geometry." A + indicates the location is reinforced, while a - indicates no reinforcement.

The M-S 2008 model was created to solve problems that emerged when the M-S 2007 model simulated the training phase of this paradigm (Dawson et al., 2008). In particular, the positiveness correction in M-S 2008 was intended to prevent the impossible probabilities generated by M-S 2007. Miller and Shettleworth (2008) found that when their revised model was used for the Wall et al. (2004) paradigm, it behaved correctly, even when high values (0.65) for $\alpha\beta$ were used. However, we show below that when some of the parameters of the revised model are slightly changed, errors are still produced. In short, the positiveness correction has not created a proper mathematical model of reorientation.

For example, when used to model the Wall et al. (2004) reorientation task at slightly higher learning-rate values than $\alpha\beta$=0.65, the M-S 2008 model predicts dramatic fluctuations in associative strength and choice probabilities when both should instead plateau, suggesting a lack of robustness in the underlying mathematics. (An analysis of the mathematical properties and problems of the M-S model is provided in a later appendix.) The fluctuations in associative strength escalate until they eventually culminate in a global divide-by-zero error. This is repeatable and can be predicted reliably; for instance, for $\alpha\beta$=0.7, the fluctuations begin between iterations 10 and 20 (depending on the cue or location) and the collapse occurs at iteration 183 (see Figure 1-3). Similar fluctuations (in addition to the singularities reported by Dawson et al. (2008) can be demonstrated in the M-S 2007 formulation as well, suggesting the fluctuation is not due to the M-S 2008 model's positiveness correction.

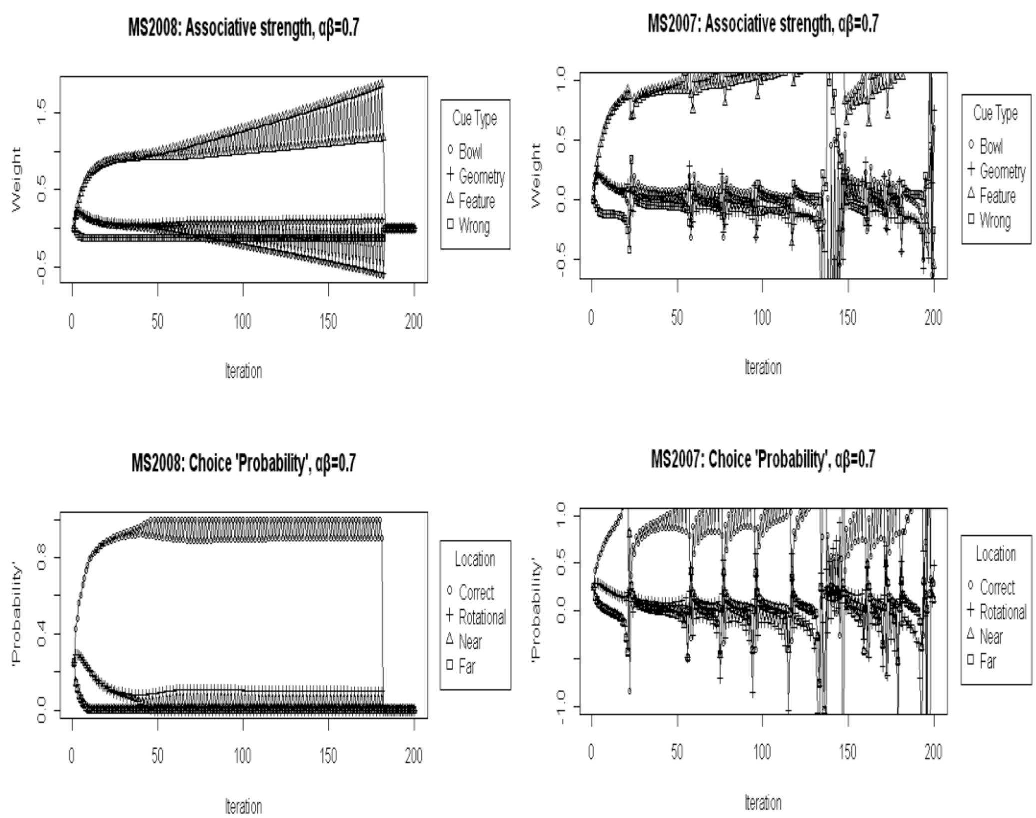

Figure 1-3. The two Miller-Shettleworth models (2007, 2008) simulating the Wall et al. (2004) reorientation task at a high $\alpha\beta$. Note the fluctuation in both models. For the M-S 2008 model (left), the 0s in the lower panel are actual 0 values, while the upper panel uses 0 as a replacement for divide-by-zero, occurring after the crash at iteration 183. The singularities present in the M-S 2007 model (right) were reported by Dawson et al. (2008), while the fluctuations are novel.

Why Does It Fail?

Dawson et al. (2008) identified a problem in the behaviour of M-S 2007, but did not explain the occurrence of this problem mathematically. Similarly, Horne and Pearce (2010) identified an empirical issue with M-S 2008, but did not attempt to explain this problem by examining the mathematics of the model. In this section, we describe how the mathematics of the model leads to a fundamental problem with how the model handles inhibition. It is this mathematical problem with handling inhibition that is the root cause of the empirical problems identified above. To begin, let us first consider the superconditioning paradigm that was illustrated in Figure 1-1.

Superconditioning

The failure to capture Horne and Pearce's (2010) superconditioning results stems from the structure of the M-S 2008 probability equation, which includes the 'positiveness correction'. This term is employed to prevent the probability of choosing a location P_L from falling below zero even when that location's net associative strength is below zero (Dawson et al., 2008). Horne and Pearce (2010) encounter such a location during stage 2 (discrimination) of their simulation: the magnitude of the negative associative strength of the consistently-inhibitory feature cue *F* far exceeds the magnitude of the positive associative strength of the other cues presented alongside it (*B* and *G*.).

During Horne and Pearce's (2010) stage 2 (discrimination), the feature cue *F* is consistently inhibited, causing its associative strength to fall. Additionally, whenever *F* is inhibited, the other cues presented with it - the context cue *B* and the correct geometry cue *G* - are also inhibited, gaining a more negative associative strength. When the total associative strength of these locations - V_L, or the sum of the associative strengths of *F*, *B*, and *G* - falls below zero, the Miller-Shettleworth probability equation (Equation 1-4) applies the positiveness correction to set the corresponding P_L to zero, and the model should subsequently predict no further changes in associative strength in response to this location. A similar story can be told for the incorrect geometry cue *W*, which is never reinforced; once the sum of associative strengths at these locations (*B*+*W*) falls below zero, the positiveness correction prevents these cues from gaining further inhibitory strength.

However, the *B* and *G* cues are still present and reinforced at other locations on the acquisition trials during this stage, allowing those cues to continue to gain associative strength at a slow rate. These cues continue to gain positive associative strength from their reinforced presentations while the Miller-Shettleworth equation artificially prevents the cues from gaining negative associative strength during their inhibited presentations as soon as the net associative strength falls below zero. Concurrently, the associative strength of these cues (*B* and *G*) grows higher than it should. Similarly, during the acquisition trials, the *B* and *G* cues continue to gain positive associative strength from presentations at the reinforced locations, while the *B*, *W* and *G* cues are all prevented from gaining adequate negative associative strength from non-reinforced locations due to the positiveness correction artificially setting such updates to 0.

During the stage 3 (experimental or control group), this effect is inflated in the experimental group relative to the control group, as the experimental group pairs the already-inflated *B* and *G* cues with a now-consistently-reinforced *F* cue (leading to higher effective P_L, and therefore a higher change in associative strength, on the correct locations), while the control group pairs these cues with a never-reinforced novel *N* cue (with 0 initial associative strength, this does not inflate the change on the correct locations relative to the incorrect ones). Accordingly, the magnitude of the changes in associative strength will be greater for experimental groups than for control groups, even on locations that are identical between the two conditions. In effect, when *F* causes *B*+*G* to become inflated during superconditioning (relative to no real change with *N*, *B*, and *G* in the corresponding control corner), the model can acquire *more* inhibition from a non-reinforced corner during superconditioning than it could from the same non-reinforced corner during control, despite both being identical cue-wise (*B*+*W*), due to Equation 1-2 being scaled by P_L.

This weakness in handling inhibition, combined with the overall scaling of all changes in associative strength, results in the model generating incorrect predictions as described by Horne and Pearce. For instance, the probability of the model selecting a correct corner during a test trial (without any feature cues) is given by the ratio between the net associative strength at that corner and the total net associative strength, or *(B+G)/((B+G)+(B+W))*. All of these cues are artificially (and identically) inflated for both groups due to the errors during the second stage (discrimination), but during the third stage (experimental or control groups), the *W* cue has associative strength of -0.35 in the experimental group, a 65% increase relative to -0.21 in the control group. Accordingly, the probability of choosing a correct corner is higher in the control group (0.94) than the experimental group (0.91).

It is perhaps unsurprising that the source of these errors is not obvious at first glance. At a fundamental level, they emerge because the model represents locations as collections of cues, where some of those cues are present at locations that are reinforced differently. This shared-cue perspective is not taken during Horne and Pearce's discussion on superconditioning. Instead, they employ the common "A+ / AX-" notation to discuss superconditioning - but in a cue competition perspective, "A" might represent some collection of cues that are not reinforced with equal likelihood. Indeed, here, "A" refers to *B+G*, which are present and reinforced in different proportions; similarly, the *W* cue (a vital part of the behaviour described above) is simply not included in "A+ / AX-" notation. As a consequence, Horne and Pearce focus on the feature cue *F* becoming a conditioned inhibitor dependent upon the context cue *B* becoming a conditioned excitor, but do not discuss that *B* is, in truth, both reinforced and not reinforced depending on the collection of cues present at a location.

Reorientation

When modelling the Wall et al. (2004) reorientation task, the Miller and Shettleworth (2008) model produces two distinct errors, both illustrated in Figure 3. The first is a series of fluctuations of both associative strengths and choice probabilities at high learning rate *αβ*; similar behaviour is observed in the Miller and Shettleworth (2007) model as well, suggesting that this problem's root cause is shared by both models. The second problem, where the equations consistently produce a divide-by-zero error after several fluctuations, is unique to the M-S 2008 model, suggesting it is a consequence of the positiveness correction. We describe the source of these problems using terminology consistent with Miller and Shettleworth's (2007) discussion of reorientation, where the task consists of four locations (Correct, Rotational, Near, and Far) that are defined in the model as combinations of four cues (Bowl (*B*), Feature (*F*), correct Geometry (*G*), and Wrong geometry (*W*)). *G* is present at the Correct and Rotational locations, while the Near and Far locations have the *W* cue. The *B* cue is present at all four locations, and the *F* cue is present only at the Correct location. All cues are initialized to 0 associative strength, except for *B*, which is initialized to 0.1 (to reflect prior experience with bowls containing reinforcement).

The fluctuations emerge from the structure of Equation 1-2, where the change in associative strength for a cue is scaled by a function of associative strength (Equation 1-3 or Equation 1-4). Both of these scaling functions can approach zero when considering locations with strong inhibitory cues, causing the model to reduce the effective change in associative strength due to lack of reinforcement relative to the effective change in associative strength due to reinforcement. That is to say, the more inhibitory a location's cues become, the less learning takes place at that location relative to others. Since these locations contain cues which are at least partially reinforced on other locations (the *B* and *G* cues in this example), these cues acquire a greater positive change in associative weight from reinforced locations and a lesser negative change in associative weight from non-reinforced locations. At low learning rates $\alpha\beta$, this artificial inflation is small relative to the net weight, and easily handled by the error-correcting Equation 1-2. However, at sufficiently[2] high $\alpha\beta$, the artificial inflation is large enough to lead to an "overcorrection" - the magnitude of the change in weights being larger than it should be.

For example, consider the *G* cue - which is initialized at 0, is presented at two locations, and is only reinforced at one. Therefore, as reported in Miller and Shettleworth, we expect its associative strength to climb (as it is reinforced), peak (since this reinforcement is not universal), and stabilize at some small positive value. With $\alpha\beta$=0.7, the peak occurs after the third iteration at V_G=0.22; the corresponding ΔV_G= -0.01 is negative, consistent with predictions - and this is

[2]What qualifies as "sufficient" varies dramatically with the structure of the problem - cue distributions, number of locations, and initial associative strengths. For the Wall et al (2004) task as described here, "sufficient" is near $\alpha\beta$=0.68.

expected to slowly decrease in magnitude with subsequent iterations. However, the subsequent iteration is not a decrease - rather, it increases, with ΔV_G=+0.003. When evaluating the change in strength for the *G* cue, the scaling nature of Equations 1-3 and 1-4 result in a lower P_L for the non-reinforced Rotational corner than for the reinforced Correct corner - and as a result, the model assigns insufficient inhibitory strength to the *G* cue. Following this slight positive increase, the model "overcorrects" with a strong negative ΔV_G to reflect the Rotational corner -but because that corner contains the *B* cue as well as the *G* cue, the negative overcorrection applies to *B* as well. On subsequent sweeps, these interlinked overcorrections produce the distinctive fluctuating behaviour seen in Figure 1-3.

This fluctuation emerges from the structure of Equation 1-2, but the eventual "crash" seen in Figure 3 results from a divide-by-zero error in Equation 1-4. This occurs because Equation 1-4 still computes the "effective" V_L for every location in the model, and then multiplies that V_L by the Heaviside step function of V_L with threshold 0. This results in V_L if V_L>0, and 0 otherwise - as locations' net associative strengths fall below zero, they are "dropped out" of the denominator of Equation 1-4. However, due to the fluctuations seen above, which grow in magnitude over time, a time will come where the magnitude of the negative associative strength of *B* and *G* together exceeds the magnitude of *F*, the only uniquely positive cue in the system. At this point, all four locations have negative net associative strength, and thus Equation 1-4 attempts to divide by zero.

If Equation 1-4 is replaced with Equation 1-3 (resulting in the M-S 2007 model), a "crash" does not occur. Instead, the model produces singularities (identified by Dawson et al., 2008, illustrated in Figure 3). Without the positiveness correction, any location's net associative strength V_L is allowed to fall below zero - or, as noted above, for *all* locations' net associative strengths to fall below zero. At this point, the denominator in Equation 1-3 flips sign from positive to negative and a singularity appears in the corresponding associative strengths and choice probabilities. As Equation 1-3 does not include an artificial substitution of 0, the divide-by-zero outcome does not happen, and if the model is allowed to continue to run, eventually the denominator of Equation 1-3 will become positive again, resulting in the next singularity in Figure 1-3, and so on.

Because these fluctuations and (in the M-S 2008 model) crashes arise after a different number of iterations depending on the chosen learning rate parameters *αβ*, and do not seem to appear during any reasonable span of time within specific ranges for *αβ*, it would appear that a necessary step in applying the Miller-Shettleworth (2007, 2008) model is missing. Specifically, one must carry out a search of parameter space to find the boundaries at which the model will fail; such a search would need to be carried out for each permutation and combination of cues and locations present within the task. However, it is informative that points of failure exist at all: such failures suggest an underlying problem with the mathematics of the model. At the end of this paper we provide a technical appendix that reveals exactly what this underlying problem is.

A Solution to the Problems

Miller and Shettleworth (2007) make an important observation concerning the application of associative models to geometry learning tasks: the agent's pattern of behaviour alters each cue's apparent reinforcement contingencies on any given trial. In short, Miller and Shettleworth note that such learning is intrinsically <u>operant</u>. Miller and Shettleworth endeavored to model this by using the P_L equation to scale the Rescorla-Wagner model as formalized in Equation 1-2. However, this scaling results in the improper handling of net negative associative weights, producing the problems that have been described above. Importantly, the choice of P_L equation - Equation 1-3 or Equation 1-4 - merely alters the form these issues take: it is within the structure of Equation 1-2 that the true problem lies. In this section, we describe an alternative model that solves these problems, but which is still both associative and operant in nature.

To begin, let us consider Miller and Shettleworth's (2007) view of operant learning as follows: at some moment in time, an agent perceives a set of cues V_L related to a particular location *L*. The agent uses these cues to make a judgment about how attractive this location is. For instance, the agent might use these cues to predict the probability P_L of being reinforced if that location is actually visited. Indeed, we could also say that P_L is the likelihood that the agent will actually visit location *L*. If the location is visited, then the agent will be reinforced (or not), and can modify the associative strengths of the available cues accordingly. Such learning is operant, because associative strengths will only be modified if the agent explores the location.

We saw earlier that one source of the problems with the M-S models is that either equation used to model P_L (i.e. either Equation 1-3 or Equation 1-4) has problems when faced with inhibition. Clearly, we need to select a different equation for P_L, one that is more robust to the negative associative strengths of inhibitory cues. A natural choice for this equation is the logistic function (given in Equation 1-5), which produces a response between 0 and 1 for all possible input values and monotonically increases as input increases.

$$\boldsymbol{P_L = \frac{1}{1 + e^{(-\Sigma V_L)}}} \tag{1-5}$$

It is impossible for the logistic function to produce a value outside of the range between 0 and 1 (solving a problem that occurs with Equation 1-3), or to result in a divide-by-zero error (solving a problem that occurs with Equation 1-4). The logistic function is an ideal choice for computing probability like P_L (Dawson & Dupuis, 2012), and has a long history of being used to model phenomena in a wide variety of disciplines (Cramer, 2003).

We also noted earlier that another source of the problems with the M-S models arose when their equations P_L were placed in the context of the remainder of Equation 1-2[3]. Importantly, the logistic equation permits us to take advantage of a different formalism that eliminates this difficulty. Equation 1-5 - which converts weighted cues into a probability - defines a modern version of a very simple artificial neural network, called a perceptron (Rosenblatt, 1958, 1962). The simplest version of a modern perceptron (Dawson, 2004, 2008) consists of a set of input units, each of which can be used to represent whether a particular cue is present or absent. Each of these input units can send a signal to a single output unit through a weighted connection; the weight of the connection represents the associative strength of a particular cue. The output unit works by summing the weighted signals from the input units to produce a single number, called net input, which is identical to $(\Sigma V_E)_L$. The output unit then produces a response - its activation - by computing the logistic function of its net input exactly as defined by Equation 1-5.

[3]For a more formal discussion of these problems and their consequences, we refer the interested reader to the technical appendix.

While Equation 1-5 defines the activation of this perceptron, it does not define the learning rule with which the connection weights are adjusted. Such a rule can be found in (Dawson, 2008), expressed here with the language of Miller and Shettleworth:

$$\boldsymbol{\Delta V_L = \eta \cdot (US_L - P_L) \cdot \alpha_E} \tag{1-6}$$

Here, η is a constant of proportionality equal to $\alpha\beta\lambda$, "US_L" reflects presence (1) or absence (0) of reinforcement (the unconditioned stimulus) at a location, a_E reflects the value of the input unit corresponding to element E, and P_L reflects the perceptron's logistic response to a given pattern of cues, given by Equation 1-5.

In employing Equations 1-5 and 1-6, we are in essence proposing that the perceptron can provide a central component of an associationist model of spatial learning. There are several reasons that this proposal is attractive. First, the mathematics of this type of model are well-established - there is a long history of mathematical results concerning perceptron learning, beginning with the work of Rosenblatt (1958, 1962). Second, the associative models of Miller and Shettleworth (2007, 2008) are extensions of the well-established Rescorla-Wagner model of associative learning (Rescorla & Wagner, 1972). Importantly, the kind of learning that is carried out by a perceptron can be formally translated into Rescorla-Wagner learning (Dawson, 2008; Gluck & Bower, 1988; Sutton & Barto, 1981). Third, the motivation behind Equation 1-5 was to generate a value that could be interpreted as a probability. It has been shown empirically that perceptron responses can be interpreted as probabilities, because these networks can learn to generate responses that match the probabilities of events occurring in the world (Dawson, Dupuis, Spetch, & Kelly, 2009). Furthermore, formal analyses of perceptrons prove that the activity of an output unit can literally be called a conditional probability (e.g. Dawson & Dupuis, 2012). Fourth, one of the reasons that the M-S models are of interest is because they have been argued to be able to model reorientation task regularities. Crucially, perceptrons have also been shown to be capable of modeling a variety of reorientation task phenomena (Dawson et al., 2010).

While the perceptron has been successfully used to model the reorientation task (Dawson et al., 2010), this was done using standard learning rules (Dawson, 2004, 2008; Rosenblatt, 1962), which are <u>not operant</u> in nature. We now describe an algorithm which trains a perceptron in an operant fashion, transporting Miller and Shettleworth's (2007) core idea about reorientation into the domain of artificial neural networks.

The typical, and non-operant, manner for training a perceptron (Dawson, 2004, 2008) proceeds as follows: First, a pattern (i.e. a set of cues, such as those corresponding to one "location") is presented to the perceptron's input units. Second, the perceptron converts input unit signals into an output response (i.e. Equation 1-5). Third, the perceptron receives feedback about its response (e.g. it receives reinforcement, or not). Fourth, a learning rule (Equation 1-6) is used to modify connection weights in accordance with the feedback.

A simple change to the above learning algorithm makes it truly operant (Dawson et al., 2009). The second step in the above procedure is to compute P_L using Equation 1-5. Once this is computed, we can add a new step where P_L is used to make a choice - in essence, a choice about whether or not to visit location L - where the likelihood of visiting the location is P_L. If the choice is made to visit the location, then learning proceeds according to the third and fourth steps in the above algorithm. However, if the choice is made to not visit the location, then no learning occurs - connection weights are not updated, and the algorithm returns to the first step when presented with another pattern. This perceptron is operant because it only learns when it chooses to act; if it does not choose to act on a given trial, its connection weights are not updated. Furthermore, it is operant in the way that Miller and Shettleworth (2007) desire, because as the associative strength of the cues at a location increase, P_L (a function of those associative strengths) increases, and so does the likelihood that location L will be visited. Conversely, as P_L decreases, so does the likelihood of visiting location L. As it learns about its environment, the operant perceptron will be more likely to choose locations (i.e. cue configurations) that lead to reinforcement, and will be less likely to choose locations that do not lead to reinforcement. As was noted earlier, and detailed in the technical appendix, the root mathematical cause of the problems with the M-S model is the fact that it scales changes in associative weights by P_L. The procedure for training an operant perceptron solves this problem by separating operant choice from weight modification. That is, Equation 1-5 is used to make a decision about whether to learn or not, and then standard learning (Equation 1-6) is conducted accordingly. Importantly, at no point in Equation 1-6 are changes in association multiplied by Equation 1-5, preventing the mathematical difficulties described in the appendix.

Previous research has shown that the operant perceptron can learn to perform a probability matching task (Dawson et al., 2009). Furthermore, this previous research has shown that the behaviour of the operant perceptron at equilibrium is similar to that of a traditional perceptron trained on the same probability matching task. This indicates that the operant training procedure does not violate the mathematical regularities associated with perceptron learning. Traditional perceptrons have been shown to be promising models of reorientation. In the next section, we demonstrate that this is also true of the operant perceptron, by showing that it generates appropriate results for both of the case studies introduced earlier in this paper.

Evaluating the Operant Perceptron

To see if the operant perceptron is capable of succeeding where the M-S model had difficulty, we simulated both the superconditioning and reorientation tasks, as described below.

Superconditioning

Method

Horne and Pearce's (2010) superconditioning experiment was presented to the operant perceptron using five inputs, corresponding to the five cues (*B*, *G*, *W*, *F*, and *N*) from Figure 1-1 above. Any cue could be presented to the perceptron by activating its input unit with a value of 1; if a cue was absent, then the activity of its input unit was 0. These cues were grouped by location and stage as described in Figure 1-1, with each training pattern (i.e. each set of available cues) representing a corner present at a given stage. The network used a single output unit with a logistic activation function with bias held constant at 0. The network's learning rate was set to 0.05, and all of its weights initialized to 0. Patterns were presented in a random order to the network

The perceptron's learning algorithm was made operant as described above. After each pattern (i.e. collection of cues at a location) was presented to the network, the perceptron computed P_L using Equation 1-5, producing a number between 0 and 1. Then, a random number between 0 and 1 was generated. If the network's activity exceeded that random number, the network was said to have "chosen" to visit the location on this trial, and its connection weights were updated. If the network's activity did not exceed this number, the network was said to have not chosen to visit the location on this trial, and no weights were updated. This process repeats for every pattern (location) present in the scenario.

Each stage received 5000 sweeps of this training before a final, geometry-only probe trial (consisting just of the *B+G* cues) was carried out. Due to the stochastic nature of the operant perceptron's training procedure (which allows identical networks to make different patterns of choices), this simulation was repeated five times for each experimental condition and the aggregate responses were averaged.

Results

After training, the experimental network had average connection weights (associative strengths) of *B*=1.3 and *G*=4.91, while the control network had connection weights of *B*=0.29, *G*=3.20. The networks' responses to these cues - the logistic function of *B+G* - are 0.998 for the experimental network and 0.970 for the control network.

Because these responses reflect the probability of investigating these correct-geometry-only locations, it can be concluded that both groups of networks were capable of learning the geometry of the task (they have a high probability of visiting a corner with only the correct geometry present). Additionally, the experimental group's probability of investigating a geometrically-correct corner is higher than that of the control group: the experimental group displays evidence of superconditioning. This is in agreement with Horne and Pearce's (2010) animal data, and distinct from their attempt to model the same task with the M-S 2008 model, which produced the opposite result (a response of 0.92 in the experimental condition, and a response of 0.95 in the control condition).

Reorientation

Our second simulation involved training the operant perceptron on the Wall et al.(2004) reorientation task. First, we modeled this task using Miller and Shettleworth's (2007, 2008) parameters, and then we tested for robustness by exploring the operant perceptron's behaviour when different parameters were used.

Method

The Wall et al. task is defined as described in Figure 1-2. The training set consists of four input units, used to represent the presence or absence of the *B*, *F*, *G*, and *W* cues. *B* was initialized to have a 0.1 initial weight, while the other three cues had 0 initial weight. The network's learning rate was set to 0.15, exactly as in Miller and Shettleworth (2007). The training procedure was altered into an operant procedure exactly as was described for the superconditioning experiment. Five such networks ran until convergence (after approximately 8500 sweeps), and their responses were averaged.

Results

The results of this simulation are displayed in Figure 1-4. These results are virtually indistinguishable from the M-S 2008 model, if that model was supplied with well-behaved parameters. It would seem that the operant perceptron model is capable of learning the reorientation task as defined by Wall et al. (2004).

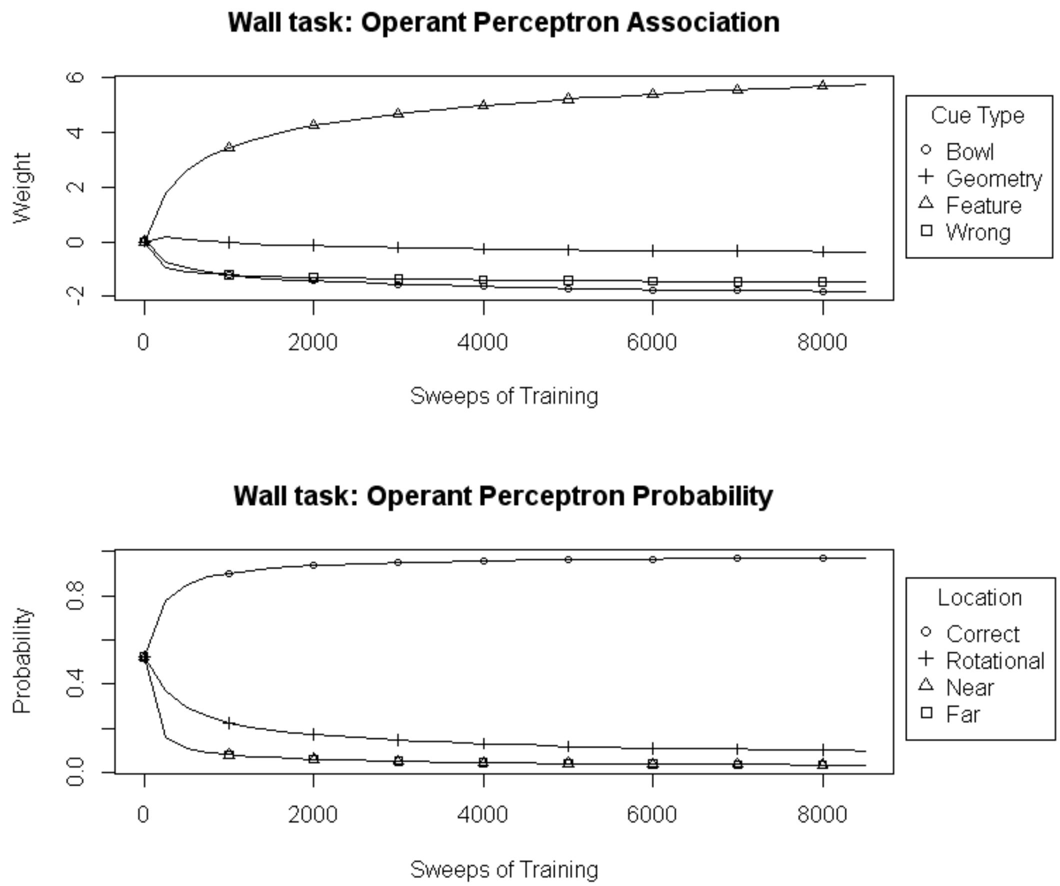

Figure 1-4: The operant perceptron's performance on the Wall et al (2004) reorientation task.

In order to test for robustness, the operant perceptron model was run again with extreme learning rates under the same initial conditions. Setting the learning rate to 0.7 did not lead to fluctuations nor to any "crash" - rather, the perceptron converges normally after approximately 1700 sweeps. Even a learning rate of 1.00 - which causes the M-S 2008 model to collapse after just 12 iterations - results in the perceptron converging in ~1300 sweeps.

In conclusion, we find that this operant perceptron model is capable of empirically handling results that the Miller and Shettleworth (2007, 2008) models are not, and that the operant perceptron's behaviour is robust to the extreme choices in learning rate that caused trouble for the M-S model.

General Discussion

Miller and Shettleworth's (2007, 2008) associative model of reorientation is rooted in the observation that reorientation is fundamentally a problem of operant learning. Therefore, they endeavored to formalize this by creating an operant version of the established Rescorla-Wagner (1972) theory of associative learning. However, their attempt to extend the Rescorla-Wagner model in this fashion has resulted in problems.

Some of these problems have already been documented in the literature. Dawson et al. (2008) discovered that the original M-S model (2007) will produce impossible probabilities under a variety of circumstances. This led Miller and Shettleworth to modify their original model (Miller & Shettleworth, 2008). However, Horne and Pearce (2010) found that this revised model did not correctly model animal data collected for tasks to which the model should apply, such as geometric superconditioning.

Other problems with the M-S model have been reported for the first time in the current paper. As shown above, when model parameters are manipulated, the behaviour of the M-S model is unstable, producing dramatic fluctuations. Indeed, for the M-S 2008 model, these fluctuations lead to an eventual "crash" that is caused when the model is required to divide a value by zero.

While previous studies (Dawson et al., 2008; Horne & Pearce, 2010) have documented some problems with the M-S models, they did not attempt to explain these difficulties. In addition to discussing some new problems, we have also shown that all of these problems emerge from the structure of Equation 1-2. In effect, the equation cannot properly handle situations with uniquely inhibitory cues when those cues are paired with other, excitatory cues. These situations lead to inappropriate scaling of changes in associative strength when Equation 1-2 is employed, and this ultimately gives rise to all of the problems described above.

However, in the current paper we have done more than demonstrate problems and trace their mathematical root. We have also provided a different model that can overcome these problems. Our alternative formalism preserves Miller and Shettleworth's (2007) operant-learning goal, but is anchored on the solid foundations of artificial neural network mathematics (e.g. Rosenblatt, 1958, 1962). We have presented simulations that show this new 'operant perceptron' model corrects these problems, and is capable of learning both reorientation and superconditioning. We believe this operant perceptron to be a plausible architecture for modeling reorientation task learning, which was the primary intent of the M-S model (Miller, 2009; Miller & Shettleworth, 2007, 2008). We have already shown that the operant perceptron can easily model other domains; for instance it can learn to match reinforcement probabilities (Dawson et al., 2009). The extent to which the operant perceptron can match the ability of the traditional perceptron to model the further intricacies of reorientation (Dawson et al., 2010) or to model a variety of classical conditioning paradigms (Dawson, 2008) is clearly a matter for future research.

References

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, *23*(2), 149–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3742991

Cheng, K., & Newcombe, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic bulletin & review*, *12*(1), 1–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15945200

Chiandetti, C., & Vallortigara, G. (2008). Spatial reorientation in large and small enclosures: comparative and developmental perspectives. *Cognitive processing*, *9*(4), 229–38. doi:10.1007/s10339-008-0202-6

Cramer, J. S. (2003). *Logit Models from Economics and Other Fields*. Cambridge, MA: Cambridge University Press.

Dawson, M. R. W. (2004). *Minds and Machines: Connectionism and Psychological Modeling*. Oxford, UK: Blackwell.

Dawson, M. R. W. (2008). *Connectionism and Classical Conditioning*. *Comparative Cognition and Behaviour Reviews* (Vol. 3, p. 115). Comparative Cognition Society. doi:10.3819/ccbr.2008.30008

Dawson, M. R. W., & Dupuis, B. (2012). The equilibria of perceptrons for simple contingency problems. *IEEE Transactions in Neural Networks and Learning Systems*, *23*(8), 1340–1344. doi:10.1109/TNNLS.2012.2199766

Dawson, M. R. W., Dupuis, B., Spetch, M. L., & Kelly, D. M. (2009). Simple artificial neural networks that match probability and exploit and explore when confronting a multiarmed bandit. *IEEE Transactions in Neural Networks*, *20*(8), 1368–1371. doi:10.1109/TNN.2009.2025588

Dawson, M. R. W., Kelly, D. M., Spetch, M. L., & Dupuis, B. (2008). Learning about environmental geometry: a flaw in Miller and Shettleworth's (2007) operant model. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(3), 415–8. doi:10.1037/0097-7403.34.3.415

Dawson, M. R. W., Kelly, D. M., Spetch, M. L., & Dupuis, B. (2010). Using perceptrons to explore the reorientation task. *Cognition*, *114*(2), 207–26. doi:10.1016/j.cognition.2009.09.006

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: an adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2971760

Horne, M. R., & Pearce, J. M. (2010). Conditioned inhibition and superconditioning in an environment with a distinctive shape. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(3), 381–94. doi:10.1037/a0017837

Miller, N. Y. (2009). Modeling the effects of enclosure size on geometry learning. *Behavioural processes*, *80*(3), 306–13. doi:10.1016/j.beproc.2008.12.011

Miller, N. Y., & Shettleworth, S. J. (2007). Learning about environmental geometry: An associative model. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(3), 191–212. doi:10.1037/0097-7403.33.3.191

Miller, N. Y., & Shettleworth, S. J. (2008). An associative model of geometry learning: a modified choice rule. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(3), 419–22. doi:10.1037/0097-7403.34.3.419

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, NY: 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*(2), 135–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7291377

Wall, P. L., Botly, L. C. P., Black, C. K., & Shettleworth, S. J. (2004). The geometric module in the rat: independence of shape and feature learning in a food finding task. *Learning & behavior*, *32*(3), 289–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15672824

Chapter 2

The Equilibria of Perceptrons for Simple Contingency Problems

Brian Dupuis and Michael R.W. Dawson

Department of Psychology, University of Alberta

Author Note:

A version of this chapter has been published. Dawson & Dupuis (2012), IEEE Transactions on Neural Networks and Learning Systems. 23(8): 1340-1344.

Introduction

A fundamental characteristic of an adaptive agent is the ability to detect causal relations (Cheng & Holyoak, 1995). However, the real world poses constant challenges to this ability, because cues do not signal outcomes with complete certainty (Dewey, 1929). It has been argued that adaptive systems deal with worldly uncertainty by becoming "intuitive statisticians", whether these systems are humans (Peterson & Beach, 1967) or animals (Gallistel, 1990; Shanks, 1995). The notion of "intuitive statistician" has been rigorously developed in a series of important papers to mean sensitivity to contingency, where contingency is defined in a normative model as a contrast between conditional probabilities (Allan, 1980; Cheng, 1997; Cheng & Holyoak, 1995; Cheng & Novick, 1990, 1992; Robert A. Rescorla, 1967, 1968). For instance, consider the simple situation that is detailed in the contingency table provided in Table 2-1. The contingency between the cue and the outcome is formally defined as the difference in conditional probabilities ΔP, where $\Delta P = P(O|C) - P(O|{\sim}C)$(Allan, 1980). More sophisticated models, such as the probabilistic contrast model (Cheng & Novick, 1990) or the power PC theory (Cheng, 1997) define more complex probabilistic contrasts that are possible when multiple cues occur, and when what they signal depends upon the context in which they are considered.

	O	**~O**
C	a	b
~C	c	d

Table 2-1: A simple contingency situation in which a cue can occur (*C*) or not (~*C*), and an outcome can occur (*O*) or not (~*O*) as well. The four letters in the table represent the frequency of co-occurrence of the two types of events. Using this table, $\Delta P = P(O|C) - P(O|{\sim}C) = A/(A + B) - C/(C + D)$.

Because many associative learning paradigms can be interpreted as teaching contingencies to humans or animals, another issue that has arisen in the literature is the relationship between formal contingency theories and formal theories of associative learning (Shanks, 2007). In particular, researchers have compared the predictions of the Rescorla-Wagner model of learning (R.A. Rescorla & Wagner, 1972) to formal theories of contingency (Chapman & Robbins, 1990; Cheng, 1997; Cheng & Holyoak, 1995). This is typically accomplished by determining equilibria for the Rescorla-Wagner model, and then comparing associative strengths of the Rescorla-Wagner model at equilibrium to probabilistic contrasts defined by contingency theory. An equilibrium of the Rescorla-Wagner model is the set of associative strengths defined by the model at the point where changes in error defined by Rescorla-Wagner learning asymptote to zero (Danks, 2003). While in some instances the Rescorla-Wagner model predicts the conditional contrasts defined by a formal contingency theory like the power PC model, in other situations it fails to generate these predictions (Cheng & Novick, 1992).

The formal results relating contingency theory to the Rescorla-Wagner model have been assumed to also apply to connectionist models of associative learning (Cheng & Holyoak, 1995; Shanks, 1995). Researchers have claimed that there is a formal equivalence (Gluck & Bower, 1988; Gluck & Myers, 2001; Sutton & Barto, 1981) between learning as defined by the Rescorla-Wagner model and learning as defined by the so-called delta rule, which is an error-correcting method that is used to train simple artificial neural networks (Dawson, 2004; Stone, 1986). Such claims are used to support the informal conclusion that any results pertaining to the relationship between the Rescorla-Wagner model and contingency theory also apply to artificial neural networks trained with the delta rule. That is, if for at least some cases *x* the Rescorla-Wagner model and contingency theory are equivalent, and if the Rescorla-Wagner model is equivalent to delta rule learning, then it seems safe to conclude that for these same cases *x* networks trained with the delta rule should be equivalent to contingency theory.

One example of this indirect argument is provided by Cheng (1992), who performs a detailed computational analysis of the relationship between the Rescorla-Wagner model and contingency theory. She emphasizes the Rescorla-Wagner model because “the learning rule it incorporates is a version of the ‘delta rule’ commonly used in connectionist models. My analysis of this model should therefore be relevant to connectionist models using this rule, whatever the content domain of the model” (Cheng & Novick, 1992, p. 371). However, Cheng neglects to conduct a computational analysis that directly relates contingency theory to artificial neural networks. Cheng and Holyoak (1995) provide a second example of the indirect argument when they use the assumed equivalence between the delta rule and the Rescorla-Wagner model to define asymptotic associative weights for the latter. A third example of this indirect argument is provided by Shanks' (1995) interpretation of the computation work of Chapman and Robbins (1990). In an appendix to their article, Chapman and Robbins prove that in a particular situation "the Rescorla-Wagner model reduces to the ΔP rule" (p. 545). However, Shanks (1995, p. 112) uses the indirect argument to interpret the proof in a connectionist light, claiming that "Chapman and Robbins have established the very important fact that the delta rule, at asymptote, yields weights that are identical to ΔP."

However, there are important reasons to be wary of using the assumed relationship between the Rescorla-Wagner model and simple artificial neural networks to infer relationships between networks and contingency theory. First, previous proofs of the formal equivalence between the Rescorla-Wagner model and the delta rule (Gluck & Bower, 1988; Gluck & Myers, 2001; Sutton & Barto, 1981) neglect to include a critical component of artificial neural networks trained by the delta rule – the nonlinear activation function that converts an output unit's net input into activation. Dawson (2008) has shown that these proofs assume a linear relationship between net input and activity, and therefore do not apply to simple neural networks such as the traditional perceptron (Rosenblatt, 1958, 1962) that uses a step function to compute output unit activity, or a modern variation of the perceptron (Dawson, 2004, 2008) that uses a logistic activation function. When the nonlinear activation function is taken into account, a formal relationship between the Rescorla-Wagner model and complete (i.e. nonlinear) networks can still be established (Dawson, 2008). However, the inclusion of the activation function imposes a crucial algorithmic difference between network learning and the Rescorla-Wagner model: the activation function serves as a theory of how internal associations are converted into network behaviour, while a theory of behaviour is not part of the Rescorla-Wagner (Miller, 2006; R.A. Rescorla & Wagner, 1972). As a result, a perceptron that uses the logistic activation can generate different behaviours than can a model trained using the Rescorla-Wagner model, and in many cases can overcome some limitations faced by the Rescorla-Wagner model (Dawson, 2008).

In short, the relationship between Rescorla-Wagner learning and artificial neural network learning is more complicated than one might expect from older comparisons (Gluck & Bower, 1988; Sutton & Barto, 1981). As a result, it is unwise to use these older analyses as the basis for an indirect link between networks and contingency theory. Instead, computational analyses that directly explore the relationships between connectionist networks and contingency theory are required. The purpose of this paper is to provide one such analysis. It is proven below that when a simple artificial neural network reaches equilibrium for a basic contingency theory problem; this equilibrium appears to be quite different from the equilibrium of the Rescorla-Wagner model for the same contingency problem. That is, in contrast to Shanks' (1995) connectionist interpretation of Chapman and Robbins' (1990) proof, the connection weights of the network are not identical to ΔP. However, ΔP can be recovered by comparing the behaviour of the network in different cue situations.

Deriving the Equilibrium

Derivation

To begin, consider the simple contingency problem that was presented earlier in Table 2-1. Chapman and Robbins (1990) proved that when Rescorla-Wagner learning reaches equilibrium for this problem the associative strength between the cue and the outcome was exactly equal to ΔP. Their proof required the assumption that there were two cues involved, the one of interest (*C*) that was present on some trials and absent on others (as in Table 2-1), and a second (*X*) that represented cues from an experimental context that were present on every trial. Rescorla-Wagner learning would alter the strengths of two associations, the one between *C* and the outcome (V_C) and the one between *X* and the outcome (V_X). For the situation defined in Table 2-1, Chapman and Robbins found that at equilibrium $V_C = a/(a + b) - c/(c + d)$. Let us now proceed to derive the equilibrium for a perceptron faced with the same contingencies.

One can train a simple perceptron on the Table 2-1 contingency problem. The perceptron would have a single input unit that would be turned on with a value of 1 when *C* is present, and turned off with a value of 0 when *C* is absent. This signal would be sent through a single connection, with connection weight w_c, to a single output unit. The desired response of this output unit would be 1 in trials in which the outcome *O* occurred, and would be 0 in trials in which *O* did not occur. On any given trial, the net input *net* (i.e. the total signal) to the output unit is equal to w_c times the activation value of the input unit. A nonlinear transformation of the net input produces the output unit's response for the trial. Let us define this nonlinear transformation as the logistic equation, which is an activation function that is commonly employed in artificial neural networks (Bechtel & Abrahamsen, 2002; Dawson, 2004; Rummelhart, Hinton, & Williams, 1986):

$$f(net) = \frac{1}{1 + e^{-(net - \theta)}} \tag{2-1}$$

The logistic equation is often described as a "squashing" function, because it is a sigmoid-shaped function that squashes values of net input, which can range from negative to positive infinity, into the range from 0 to 1. In Equation 2-1, *net* is the net input from the perceptron's input unit, and θ is a constant that is called the bias of the logistic equation. When net input equals θ, the logistic equation returns a value of 0.5. θ can be described as the value of a weight between an "extra" input unit and the output unit, where this "extra" input unit has an activation value of 1 for every pattern that the perceptron is presented. In other words, the use of θ in the logistic equation is equivalent to Chapman and Robbins' (1990) use of an extra cue to represent the constant presence of experimental context (Dawson, 2008).

Assume that when the cue is present, the logistic activation function computes an activation value that we will designate as o_c, and that when the cue is absent it returns the activation value designated as $o_{\sim c}$. We can now define the total error of responding for the perceptron (i.e. its total error for the ($a + b + c + d$) number of patterns that represent a single "sweep" in which each instance of the contingency problem given in Table 2-1 is presented once). For instance, on a trial in which *C* and *O* both occur (i.e. both *C* and *O* equal 1) the perceptron's error for that trial is the squared difference between *O* and o_c. As there are *a* of these trials, the total contribution of this type of trial to overall error is $a(1 - o_c)^2$. Applying this logic to the other three cells of Table 2-1 overall error *E* can be defined as follows:

$$E = a(1 - o_c)^2 + b(0 - o_c)^2 + c(1 - o_{\sim c})^2 + d(0 - o_{\sim c})^2 \tag{2-2}$$

$$= a(1 - o_c)^2 + b(o_c)^2 + c(1 - o_{\sim c})^2 + d(o_{\sim c})^2$$

For a perceptron to be at equilibrium, it must have reached a state in which the error term defined in Equation 2-2 has been optimized, so that error can no longer be decreased by using the learning rule to alter the perceptron's weight. To determine the equilibrium of the perceptron for the Table 2-1 problem, we begin by taking the derivative of Equation 2-2 with respect to the activity of the perceptron when the cue is present (o_c). This derivative is presented as Equation 2-3. We also need to determine the derivative of Equation 2-2 with respect to the activity of the perceptron when the cue is not present ($o_{\sim c}$). This derivative is presented as Equation 2-4.

$$\frac{\partial E}{\partial o_c} = 2(a(o_c - 1) + \ bo_c) \tag{2-3}$$

$$\frac{\partial E}{\partial o_{\sim c}} = 2(c(o_{\sim c} - 1) + do_{\sim c}) \tag{2-4}$$

One condition of the perceptron at equilibrium is that o_c is a value that causes the derivative in Equation 2-3 to be equal to 0. In Equation 2-5 this derivative is set to 0 and the equation is solved to determine the value of o_c. The reader will note that this value is equal to *a/(a + b)*, which is equal to the conditional probability *P(O|C)*.

$$0 = 2(a(o_c - 1) + bo_c) \tag{2-5}$$

$$= a(o_c - 1) + bo_c$$

$$= \ ao_c - \ a + bo_c$$

$$a = \ o_c(a + b)$$

$$\frac{a}{a + b} = \ o_c$$

$$P(O|C) = o_c$$

A second condition of the perceptron at equilibrium is that $o_{\sim c}$ is a value that causes the derivative in Equation 2-4to be equal to 0. In Equation 2-6 this derivative is set to 0 and the equation is solved to determine the value of $o_{\sim c}$. The reader will note that this value is equal to *c/(c + d)*, which is equal to the conditional probability *P(O|~C)*.

$$\mathbf{0} = \mathbf{2}(\boldsymbol{c}(\boldsymbol{o}_{\sim\boldsymbol{c}} - \mathbf{1}) + \boldsymbol{d}\boldsymbol{o}_{\sim\boldsymbol{c}}) \tag{2-6}$$

$$= \boldsymbol{c}(\boldsymbol{o}_{\sim\boldsymbol{c}} - \mathbf{1}) + \boldsymbol{d}\boldsymbol{o}_{\sim\boldsymbol{c}}$$

$$= \boldsymbol{c}\boldsymbol{o}_{\sim\boldsymbol{c}} - \boldsymbol{c} + \boldsymbol{d}\boldsymbol{o}_{\sim\boldsymbol{c}}$$

$$\boldsymbol{c} = \boldsymbol{o}_{\sim\boldsymbol{c}}(\boldsymbol{c} + \boldsymbol{d})$$

$$\frac{\boldsymbol{c}}{\boldsymbol{c} + \boldsymbol{d}} = \boldsymbol{o}_{\sim\boldsymbol{c}}$$

$$\boldsymbol{P}(\boldsymbol{O}|\sim\boldsymbol{C}) = \boldsymbol{o}_{\sim\boldsymbol{c}}$$

To provide a concrete example of the implications of these equations, let us consider the result of training a perceptron on a 'toy problem' consistent with Table 2-1. Imagine a training set consisting of 20 patterns, each involving a single cue represented by the activation of a perceptron that has only one input unit. The cue is present in exactly half of these patterns, and is reinforced (i.e. the perceptron is trained to output a value of 1.0) for 8 of these training patterns, and is not reinforced (i.e. the perceptron is trained to output a value of 0.0) for the remaining 2 patterns. The cue is absent in the remaining 10 patterns, 2 of which are reinforced, while the remaining 8 are not reinforced. This statement of the problem permits the four entries of Table 2-1 to be filled out as follows: a = 8, b = 2, c = 2, and d = 8. For these table values, $\Delta P = (a/(a + b)) – (c/(c + d))$ = (8/(8+2)) –(2/(2 + 8)) = 0.6. Using software developed in our lab (Dawson, 2005), a gradient descent rule was used to train a perceptron on this problem using a learning rate of 0.1, with the bias of the output unit and the connection weight randomly initiated in the range [-0.1, 0.1]. 400 training epochs, in which each of the 20 patterns is presented once in random order, were conducted; after 400 epochs the network had stabilized. At the end of this training, the weight of the connection between the input unit and the output unit was 2.76, and the bias of the output unit was -1.38. When the cue was presented by turning the input unit on, an output value of 0.8 was generated, which is $P(O|C)$. When the cue was not presented by turning the input unit off, an output value of 0.2 was presented, which is $P(O|{\sim}C)$.

Implications

One implication of the proof developed above is that for the type of contingency problem described in Table 2-1, at equilibrium the output of a perceptron trained on this problem can literally be described as a conditional probability. When the cue is present, perceptron output can be literally interpreted as the likelihood of the outcome given the cue. Similarly, when the cue is absent, perceptron output can be literally interpreted as the likelihood of the outcome in the absence of the cue. This was shown in the toy example provided above, where the perceptron activity was equal to the appropriate conditional probability depending upon the presence or absence of the cue.

This result makes contact with the extensive empirical literature on probability matching. Probability matching occurs when the probability with which an agent makes a choice among alternatives mirrors the probability associated with the outcome or reward of that choice (Vulkan, 2000). Studies involving a variety of subjects, including insects, fish, turtles, pigeons, and humans have not only shown the existence of probability matching, but have also shown that probability matching is adaptive: when the probability of reinforcement associated with a cue changes, the choice probabilities exhibited by the agent are quickly adjusted (Behrend & Bitterman, 1961; Estes & Straughan, 1954; Fischer, Couvillon, & Bitterman, 1993; Graf, Bullock, & Bitterman, 1964; Keasar, Rashkovich, Cohen, & Shmida, 2002; Kirk & Bitterman, 1965; Longo, 1964; Niv, Joel, Meilijson, & Ruppin, 2002). It was recently shown that perceptrons that use the logistic activation function match probabilities, and also quickly adapt these probabilities when reinforcement contingencies are altered (Dawson, Dupuis, Spetch, & Kelly, 2009). The proof above grounds this empirical finding in mathematics by demonstrating that perceptron outputs are identical to conditional probabilities.

A second implication of the proof developed above is that an equilibrium for a perceptron faced with the Table 2-1 contingency problem is not, as expected by Shanks (1995), identical to the equilibrium for the Rescorla-Wagner model. At equilibrium, the associative strength for the cue C that is determined by Rescorla-Wagner training is literally ΔP. This is not the case for the perceptron. This was shown, for instance, in the example given above; in the network that was trained neither the connection weight nor the bias was equal to ΔP.

Importantly, the fact that the associative strengths at equilibrium for the Rescorla-Wagner model differ from those at equilibrium for the perceptron does not indicate qualitative differences between the two in the context of the contingency problem being solved. That is, the two systems achieve equilibria that appear to be different because the two systems use associative strengths in different ways to produce behaviour (i.e. to generate judgments of contingency). For the Rescorla-Wagner model, the general assumption is that associative strengths are converted into responses by a linear transformation (Dawson, 2008). Thus, if the behaviour of such a model is to reflect ΔP, then ΔP must be directly represented in associative strengths, as proved by Chapman and Robbins (1990). In contrast, the perceptron uses a nonlinear transformation when it converts associative strengths into responses. Therefore ΔP cannot be directly encoded as a connection weight. Instead, ΔP must be computed after a response is generated -- by taking the difference between a perceptron's output when the cue is present and its output when the cue is absent. For instance, in the example provided earlier, if after training one takes the difference between perceptron activity when the cue is present (0.8) and perceptron activity when the cue is absent (0.2), the result is 0.6, which is the value of ΔP given the representation of that problem in Table 2-1 format.

It might be argued that a proper difference between the two equilibria has not been established because one is framed in terms of associative strength, while the other is framed in terms of perceptron output. However, the value of θ and the value of the connection weight w_c can easily be computed given the results in Equations 2-5 and 2-6. First, if one sets the value of the logistic function in Equation 2-1 to $c/(c + d)$, assumes $net = 0$, and solves for θ, then it is found that θ equals $ln(d/c)$. Second, if one sets the value Equation 2-1 to $a/(a + b)$, assumes $\theta = ln(d/c)$, and solves for w_c, then it is found that w_c equals $ln(d/c) - ln(b/a)$ -- which is not equal to ΔP. (One can solve for w_c in this case because in this simple network, when $C = 1$, $net = w_c$.)

A third implication of the proof developed above is that one cannot naively assume that the formal equivalence of Rescorla-Wagner learning and delta rule learning (Gluck & Bower, 1988; Gluck & Myers, 2001; Sutton & Barto, 1981) also establishes that the Rescorla-Wagner model is identical to a connectionist network like the perceptron. The analysis of the perceptron's equilibrium reveals a final state that is structurally quite different from that predicted from Shanks'

(1995) interpretation of the Chapman and Robbins (1990) proof. That is, for the perceptron, ΔP is not directly represented as a connection weight.

This simply suggests that further formal research is required to directly establish the relationship between contingency theory and artificial neural networks. Modern contingency theory is concerned with contrasts between probabilities in situations involving multiple cues, and Danks (2003) has demonstrated how equilibria for Rescorla-Wagner models can be computed in multiple-cue situations. Future formal research is required to determine equilibria for artificial neural networks in multiple-cue situations in order to investigate the degree of agreement or disagreement between networks and contingency theory. Beginning such work with the study of simple perceptrons is likely to bear fruit, because these simple networks are still the source of surprising and interesting results (Fernandez-Delgado, Ribeiro, Cernadas, & Ameneiro, 2011; Raudys, Kybartas, & Zavadskas, 2010), and because the behaviour of perceptrons in multiple cue situations suggests that this simple kind of network can mimic core empirical regularities. For instance, one key aspect of adaptive animal behaviour is using multiple cues to maximize survival, and to use changes in the information provided by multiple cues to modify behaviour accordingly (Gallistel, 1990). Perceptrons have been shown to demonstrate such abilities, for instance by reacting to new combinations of multiple cues to modify response probabilities in a navigation task (Dawson, Kelly, Spetch, & Dupuis, 2010). It would be expected that formal analyses of the equilibria of such networks would shed a great deal of insight about their relation to more sophisticated versions of contingency theory.

A fourth implication of our results follows from the third: if naïve assumptions about the equivalence between the Rescorla-Wagner and neural networks are incorrect (as we have demonstrated), then a more rigorous account of the relationship is likely to shed new insights into the relationships between Rescorla-Wagner learning, neural network models, and contingency theory. In particular, mathematical knowledge concerning neural networks may provide new approaches to understanding learning about contingency.

For example, the proof developed above was based on a quadratic definition of network error, because this formulation of error has been central to studying the relation between Rescorla-

Wagner and neural network learning (Dawson, 2008; Gluck & Bower, 1988; Sutton & Barto, 1981). However, other definitions of error are possible (Raudys et al., 2010). For instance, some researchers have suggested that network error for noisy or stochastic environments might be better characterized in terms of measures of entropy (Baum & Wilczek, 1988; Hopfield, 1987; Solla, Levin, & Fleisher, 1988; Wittner & Denker, 1988), or equivalently using error metrics that maximize information (Plumbley, 1996, 1997, 1999). Future research that explores the relationships between contingency theory, animal learning, and neural networks using the mathematics of information theory is likely to produce interesting and important results.

References

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*(3), 147–149.

Baum, E. B., &Wilczek, F. (1988).Supervised learning of probability distributions by neural networks.In D. Z. Anderson (Ed.), *Neural Information Processing Systems* (pp. 52–61). New York, NY: AIP.

Bechtel, W., & Abrahamsen, A. A. (2002).*Connectionism and the Mind: Parallel Processing, Dynamics, and Evolution in Networks* (2nd ed.). Malden, MA: Blackwell.

Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & cognition*, *18*(5), 537–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2233266

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405. doi:10.1037//0033-295X.104.2.367

Cheng, P. W., &Holyoak, K. J. (1995). Complex Adaptive Systems as Intuitive Statisticians: Causality, Contingency, and Prediction. In H. L. Roitblat& J.-A. Meyer (Eds.), *Comparative Approaches To Cognitive Science* (pp. 271–302). Cambridge, MA: MIT Press.

Cheng, P. W., &Novick, L. R. (1990).A probabilistic contrast model of causal induction.*Journal of personality and social psychology*, *58*, 545–567.

Cheng, P. W., &Novick, L. R. (1992).Covariation in Natural Causal Induction.*Psychological Review*, *99*(2), 365–382.

Danks, D. (2003). Equilibria of the Rescorla–Wagner model.*Journal of Mathematical Psychology*, *47*(2), 109–121. doi:10.1016/S0022-2496(02)00016-0

Dawson, M. R. W. (2004). *Minds and Machines: Connectionism and Psychological Modeling*. Oxford, UK: Blackwell.

Dawson, M. R. W. (2005). *Connectionism: A Hands-On Approach*. Oxford, UK: Blackwell.

Dawson, M. R. W. (2008). *Connectionism and Classical Conditioning*.*Comparative Cognition and Behaviour Reviews* (Vol. 3, p. 115).Comparative Cognition Society. doi:10.3819/ccbr.2008.30008

Dawson, M. R. W., Kelly, D. M., Spetch, M. L., & Dupuis, B. (2010).Using perceptrons to explore the reorientation task.*Cognition*, *114*(2), 207–26. doi:10.1016/j.cognition.2009.09.006

Dewey, J. (1929). *Experience and Nature* (2nd ed.). Chicago, IL: Open Court Publishing.

Fernandez-Delgado, M., Ribeiro, J., Cernadas, E., &Ameneiro, S. B. (2011). Direct parallel perceptrons (DPPs): Fast analytical calculation of the parallel perceptrons weights with margin control for classification tasks. *IEEE Transactions in Neural Networks*, *22*(11), 1837–1848.

Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, MA: MIT Press.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: an adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2971760

Gluck, M. A., & Myers, C. (2001).*Gateway to Memory: An Introduction to Neural Network Modeling of the Hippocampus and Learning*. Cambridge, MA: MIT Press.

Hopfield, J. J. (1987). Learning algorithms and probability distributions in feed-forward and feed-back networks.*Proceedings of the National Academy of Sciences of the United States of America*, *84*(23), 8429–33. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=299557&tool=pmcentrez&rendertype=abstract

Miller, R. R. (2006). Challenges Facing Contemporary Associative Approaches to Acquired Behaviour.*Comparative cognition & behaviour reviews*, *1*, 77–93. doi:10.3819/ccbr.2008.10005

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological bulletin*, *68*(1), 29–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6046307

Plumbley, M. D. (1996). Information processing in negative feedback neural networks.*Network (Bristol, England)*, *7*(2), 301–5. doi:10.1088/0954-898X/7/2/010

Plumbley, M. D. (1997). Information theoretic approaches to neural network learning.In A. Browne (Ed.), *Neural Network Perspectives on Cognition and Adaptive Robotics* (pp. 72–90). Bristol, UK: Institute Physics Publishing.

Plumbley, M. D. (1999). Do cortical maps adapt to optimize information density? *Network (Bristol, England)*, *10*(1), 41–58. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10372761

Raudys, R., Kybartas, R., &Zavadskas, E. K. (2010).Multicategory nets of single-layer perceptrons: complexity and sample-size issues. *IEEE Transactions in Neural Networks*, *114*(2), 207–226.

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, NY: Appleton-Century-Crofts.

Rescorla, Robert A. (1967). Pavlovian conditioning and its proper control procedures.*Psychological review*, *74*(1), 71–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5341445

Rescorla, Robert A. (1968). Probability of shock in the presence and absence of CS in fear conditioning.*Journal of comparative and physiological psychology*, *66*(1), 1–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5672628

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.

Rummelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors.*Nature*, *323*, 533–536.

Shanks, D. R. (1995).*The Psychology of Associative Learning*. Cambridge, UK: Cambridge University Press.

Shanks, D. R. (2007).Associationism and cognition: human contingency learning at 25. *Quarterly journal of experimental psychology (2006)*, *60*(3), 291–309.doi:10.1080/17470210601000581

Solla, S. A., Levin, E., & Fleisher, M. (1988). Accelerated Learning In Layered Neural Networks. *Complex Systems*, *2*, 625–639.

Stone, G. O. (1986). An analysis of the delta rule and the learning of statistical associations. In D. E. Rummelhart& J. McClelland (Eds.), *Parallel Distributed Processing* (pp. 444–459). Cambridge, MA: MIT Press.

Sutton, R. S., &Barto, a G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological review*, *88*(2), 135–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7291377

Wittner, B. S., &Denker, J. S. (1988). Strategies for teaching layered networks classification tasks. In D. Z. Anderson (Ed.), *Neural Information Processing Systems* (pp. 850–859). New York, NY: AIP.

Chapter 3

fAARS-Lite: An Open Platform for Investigating Spatial Tasks in Humans

Brian Dupuis[1], Danielle M. Lubyk[1], Lucio A. Gutiérrez[2],
Marcia L. Spetch[1],& Michael R.W. Dawson[1]

[1:] Department of Psychology, University of Alberta
[2:] Department of Computer Science, University of Alberta

Author Note:

[This chapter is a modified form of a manuscript being prepared for submission by Dupuis, Lubyk, Gutiérrez, Spetch, and Dawson.]

Introduction

Empirical regularities in behaviour across species can aid in developing animal models of human cognition. In particular, comparative spatial cognition – the study of how human and non-human animals orient and find their way in their environments – has a long history of benefiting from comparative methods. For example, Tolman's (1948) classic maze experiments with rats included a mental map model of human spatial navigation, and directly led to O'Keefe and Nadel's (1978) influential cognitive mapping theory. The diverse array of species used to model spatial cognition (e.g., domestic chicks, pigeons, rats) has given rise to an impressive array of paradigms to employ with human subjects. (For a more complete overview of comparative experimental spatial paradigms, see Shettleworth, 2010).

In order to evaluate the relevance of experimental data collected from nonhuman species, scientists perform analogous experiments using human subjects. However, there are different logistical and ethical constraints on carrying out animal experiments with human subjects, the most dramatic of which is representing the precise environments used in animal experiments in a form conductive to human research. Numerous methods exist for bridging this difference. Collecting data on human performance in spatial awareness and navigation tasks typically involves an artificial, tightly-controlled environment. In one extreme, these environments may be immersive, including exact physical rooms (e.g. Newcombe, Ratliff, Shallcross, & Twyman, 2010), virtual reality technology (e.g. Zhao, Zhou, Mou, Hayward, & Owen, 2007), or even "augmented reality" systems merging the real and virtual worlds (e.g. Mou, Biocca, Owen, Tang, & Lim, 2004). However, the use of immersive environments often requires substantial space, expensive equipment, or both. As these constraints further limit the number of simultaneous participants, these methods also require extended data collection periods.

As an alternative, some research paradigms are implemented in a non-immersive manner, typically on readily-available desktop computers. These can range from static images (Kelly & Bischof, 2005) to customized commercial first-person video games (Sturz & Bodily, 2010; Talbot, Legge, Bulitko, & Spetch, 2009). Non-immersive methods are simpler and less expensive to deploy than virtual reality equipment, and can run multiple subjects simultaneously with ease, allowing for faster data collection. Furthermore, visual spatial learning experiments performed in appropriate non-immersive virtual environments appear to produce results indistinguishable from those performed in real-world versions of the same task (Kelly & Gibson, 2007; Sturz, Bodily, Katz, & Kelly, 2009), suggesting that non-immersion is not a drawback for these purposes.

However, these non-immersive methods often create their own barriers to research: restrictive software licenses and unique internal programming languages may present obstacles to researchers (especially those without extensive programming expertise), and as a result involve significant effort spent on merely implementing new experiments instead of actually running them. For instance, designing new environments in the popular Half-Life 2 game platform (Valve Software, Bellevue, WA) requires sophisticated knowledge of C++ in order to script basic events, and the terms of its academic license prevent sharing of specific methods between institutions. To summarize, while non-immersive three-dimensional platforms may provide similar results to their immersive counterparts in spatial tasks, even setting up these programs can be time consuming and expensive, often requiring advanced programming knowledge. What is needed is a user-friendly platform which does not require advanced programming skills and is easy to set up and deploy.

This paper presents an alternative to traditional game platforms: a non-immersive virtual world and experimental engine constructed using entirely open source tools available under free software licenses[4], designed to be extended to a wide array of spatial navigation or orientation experiments, and deployable with a minimum level of technical complexity. Users with no prior programming knowledge are able to set up and run classic spatial experiments with relative ease, and more experienced technicians can extend the logic to handle non-standard tasks far beyond those presented here. This virtual world is a simplified deployment of the fAARS system (for Augmented Alternate Reality Services, Gutiérrez, 2012), and as such is referred to as “fAARS-Lite” here. Here, we describe the structure of this system and provide examples of its use with basic spatial navigation paradigms.

[4]For more information on free software licenses, including details on those used here, visit the Free Software Foundation at http://www.gnu.org/licenses/license-list.html.

Generally speaking, with no modification, fAARS-Lite is capable of modeling any experiment in which subjects move freely through a specified environment, reach a particular location, (optionally) receive feedback, and then are made to repeat the process, whether in the same environment, an altered environment, or a novel environment. No special interaction with the world beyond movement is needed from subjects by default (i.e. subjects need only move to a location, as opposed to clicking on virtual objects). Several classic spatial navigation paradigms fit this template, including the radial arm maze (Olton & Samuelson, 1976), the Morris water maze (Morris, 1984), and the geometric arena (Cheng, 1986). Details on implementing these paradigms using fAARS-Lite are included below.

Apparatus Structure

fAARS-Lite consists of three broad systems: an open-source client-server virtual world package, a collection of trigger scripts placed within this world, and an external control engine that coordinates the experiment based on the simulation's current state and user-specified rules. As a system, it can be considered an event-driven machine, existing in a specific combination of states (locations, rules, conditions, and so on) until particular events (subject actions) occur to change the system. A schematic representation of these systems and the relationships between them is given in Figure 3-1.

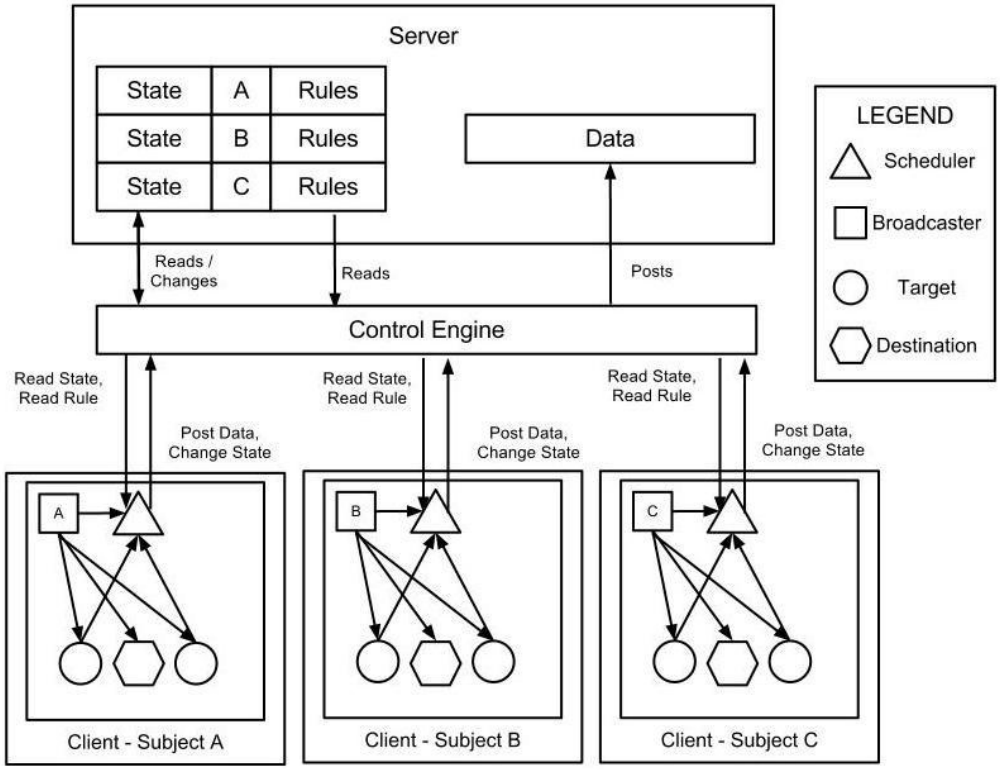

Figure 3-1: The fAARS-Lite system schematic. Not pictured: Signal objects, which would appear or disappear over each client's interface in response to Target activity.

Virtual World

The virtual world has been implemented in the client/server architecture style, allowing for multiple subjects to participate in experiments simultaneously. The communication between the server and the client is done asynchronously through HTTP using a set of RESTful APIs.

The virtual world client is constructed using the OpenSimulator ("OpenSim") program (http://www.opensimulator.org), an open-source alternative to the commercial Second Life platform (Linden Labs, San Francisco, CA.). A copy of the virtual world is installed and run locally on each client computer, along with a viewer program to display the environment in a form recognizable to the subject. We used the Hippo Viewer (MJM Labs, http://mjm-labs.com/viewer/) for this task, as it is also freely available under an open license. OpenSim and Hippo Viewer communicate through HTTP using the XML-RPC method. Together, they allow the researcher to construct a three-dimensional virtual world, known as a *region*, using the mouse and a graphical interface. By default, these clients include a visible third-person avatar; in our distribution, we render the avatar invisible, which produces the subjective effect of first-person movement.

The virtual world server consists of a common database running on a MySQL server (MySQL AB, http://www.mysql.com). This database serves two purposes: maintaining a log of all observations sent to the server from the client, and containing the current state of every active experiment. These states consist of valid sets of stimuli (e.g., a training set and a testing set) and a list of rules for moving between these states, both specified by the researcher. These states and rules are provided to the server in the form of a CSV file generated in any spreadsheet package (such as Microsoft Excel or OpenOffice).

Trigger Scripts

fAARS-Lite contains five basic classes of control objects used to direct the flow of each experiment in response to the subject's behaviour. These objects contain prepared scripts written in OpenSim's internal LLScript language. Although they can be extended, they are capable of replicating many classes of spatial task without further modification:

1. *Targets*, typically attached to a goal region or action trigger, are locations of interest for a subject, e.g., corners or specific locations in arenas that the subjects can select.
2. *Destinations*, used to mark different start locations or experimental conditions.
3. *Signals*, which temporarily override the subject's vision, and allow messages or selection feedback to be sent while hiding certain effects from view.
4. The *Scheduler* serves as a controller and communicator, linking all the components of the client's virtual world with information supplied by the server.
5. The *Broadcaster* broadcasts specific parameters throughout the virtual world, allowing a researcher to quickly configure the system for a new participant or experiment from within the virtual world itself.

A typical virtual world contains one scheduler, one broadcaster, and any number of signals, destinations, or targets as appropriate. Other objects besides these are also possible but need not contain any additional scripts; for example, inert walls may be created within Hippo Viewer to restrict a subject's movement or vision, or distinctive landmarks added to provide feature stimuli. These objects are placed in the region graphically, within Hippo Viewer, using the mouse to drag-and-drop them from a menu. Beyond filling in a unique ID name for each object, further alteration of the trigger scripts is not required.

Unless modified, each Target will wait until the subject approaches it, at which point it will activate a Signal and send data to the control engine. The Scheduler will receive new instructions from the engine based on this data, and will move the subject to the next Destination, facing a random direction. The Broadcaster is used during setup, providing global variables such as a unique subject ID and the address of the server to the other objects.

Control Engine

fAARS-Lite uses the same engine as the full fAARS package (Gutiérrez, 2012). This engine is in charge of processing each client's (i.e., participant's) rules and states according to the logic of the experiments stored in a database. It has been implemented as an event-driven service-oriented architecture using PHP, managed using the open-source program PHPMyAdmin (http://www.phpmyadmin.net). It runs on the Apache Web Server (http://www.apache.org), which allows us to expose the engine's functions to the clients as a set of RESTful APIs.

The machine consumes events that happen in the client side - events such as a subject moving onto a Target. In response to this event, the engine reads and applies the appropriate rule from the server's database, and changes the subject's current state accordingly (e.g., providing the participant with feedback when it makes a choice). Once the rules have been applied, and the state of the subject is updated and saved in the server's database, the engine sends a signal to the corresponding client, allowing the client to re-locate the subject in the virtual world according to their state during the experiment. All of these procedures are hidden from the subject through behind-the-scenes operations and use of virtual world elements (i.e. Signal objects).

Examples of Implementation

Within this general structure, experimenters can easily implement classic animal spatial paradigms with human participants using fAARS-Lite. Here, we consider three common spatial experiments: the radial arm maze (Olton & Samuelson, 1976), the Morris water maze (Morris, 1984), and the geometric arena (Cheng, 1986) and discuss their general implementation using fAARS-Lite. A basic outline for building each required virtual environment (placing inert objects and trigger scripts within Hippo Viewer using the mouse) and defining the experimental protocol (the CSV spreadsheet containing the set of states and rules uploaded to the server) is provided for each experimental paradigm.

Radial arm maze

In the simplest form of this paradigm (Olton & Samuelson, 1976), subjects are introduced into a hub surrounded by a number of arm-like hallways which are, in principle, indistinguishable from each other from the perspective of the subject apart from specific cues supplied by the experimenter. In the original study, these cues involved “extramaze” information, such as the furniture surrounding the maze, but in practice these can vary. In the basic (free-choice) condition, the subject is allowed to freely wander through the halls until it finds reinforcement in the form of food at the end of one of these arms. On subsequent trials, these arms may be rotated or shuffled (e.g., changing the relative position of the halls to stable cues, to each other, or to both) to test which cues the subject used in isolating the reinforced arms from the non-reinforced arms. The radial arm maze was originally conducted with rats (Cole & Chappell-Stephenson, 2003; Olton & Samuelson, 1976), but has since been modified and conducted with pigeons (DiGian & Zentall, 2007; Spetch & Edwards, 1986), bees (Brown & Demas, 1994), gerbils (Wilkie & Slobin, 1983), and gorillas (MacDonald, 1994). Analogues of Olton and Samuelson’s radial arm maze have been carried out with human participants, with varying results. However, practical limitations (e.g., size) of replicating a similar real-world task with humans have restricted these experiments to extremely small-scale tasks where participants are required to open the lids on containers which are arranged in a circular pattern on a table (Abrahams, Pickering, Polkey, & Morris, 1997; Rahman, Abrahams, & Jussab, 2005) or work with a drawn-on-paper version of the task (O’Connor & Glassman, 1993). While these analogues may be similar in design, both the scale of the experimental set-up and the lack of human immersion (i.e., first-person view from within the radial arm maze) present potential limitations.

Implementing the classical radial arm maze (Olton & Samuelson, 1976) in fAARS-Lite and allowing research with human participants in a virtual-world setting would be straightforward.

The region consists of walls defining eight assorted hallways; these walls do not extend above the subject's field of view, allowing for a fixed set of extramaze information (additional graphical elements) to be included in the region. A Destination object is placed in the center of the hub, and a uniquely-named Target is placed at each end. This enclosure, along with all extramaze information, is copied for each of the different rotations or shufflings planned for the experiment; these separate shufflings can be masked from the subject through large dividing walls placed beyond the extramaze cues. The Scheduler and Broadcaster are placed outside the main enclosures, out of view and accessible only by the experimenter.

On the server, we upload a CSV representing the basic protocol. That CSV might broadly define two "phases" - free-choice, and test. Free-choice would define some number of the Targets as reinforced (all other Targets are non-reinforced), and would require the subject to make a certain number of choices, receiving feedback from a Signal each time, before advancing to testing. The test phase takes place in a separate copy of the arena, with the arms shuffled and reinforcement conditions changed.

A participant would experience this study as if he were moving through a radial arm maze. When the participant reaches the end of a hallway, he/she is informed whether the choice was correct or incorrect via a splash screen (a Signal object which appears in response to a subject touching a Target). This would remain visible for a few seconds, completely obscuring the participant's view of the maze. By the time this screen disappears, the participant finds him/herself back in the center of the arena, facing a random direction, and the experiment continues. During the shift from "free-choice" to "testing", nothing subjective would change: the feedback Signal splash screen fades and the participant finds himself in essentially the same arena, except now the hallways that corresponded to earlier reinforcement may be in different directions. During this time, the server automatically logs every Target the participant walks into in the order they are chosen. This allows the researcher to check for any systematic search patterns predicted by the specific cues present in this task.

The Morris water maze

In this classic paradigm (Morris, 1984), subjects (typically rats) are placed in a geometric aquarium filled with an opaque water mixture. During training, they are led to a submerged platform by way of a visible beacon; presumably they also acquire information about the location of the platform from the shape of the arena during this time. During testing, the beacon is removed, and the subject's search pattern is recorded. Although designed for rats, human versions of the Morris water maze task have been performed, employing direct physical analogues (Newman & Kaszniak, 2010), customized software (Hamilton, Driscoll, & Sutherland, 2002), and commercial software specially adapted to the task (Skelton, Ross, Nerad, & Livingstone, 2006).

The fAARS-Lite interpretation of the Morris water maze is straightforward, even though the subjects are not "swimming". The region consists of a large geometric arena with a Destination placed in the center, with a single invisible Target placed at some location to represent the submerged platform. Two copies of the arena exist - a training arena with an extra visible beacon (e.g. a pillar) over the Target, and a testing arena where this beacon is absent.

The server-side control CSV for this basic task is essentially identical to that of the radial arm task described above - subjects are allowed to roam freely until they find the Target, whereupon they are reinforced and returned to the training environment's Destination, facing a random direction. After completing a specified number of training trials, the server deposits them in the testing arena, and subjects must find the Target without relying on the visible beacon.

The basic form of data from a water maze task involves viewing a recording of the animal and scoring its time spent searching for the platform in particular regions. This can be accomplished directly within fAARS-Lite through a screen-capture program such as Fraps (http://www.fraps.com), but it also serves as an example of how a technician can modify the fAARS-Lite logic for specific tasks. By default, the Target script records the subject's choice, displays a feedback Signal screen, and teleports the subject to the next Destination the server provides. It is trivial to remove these last two functions, creating a "Dummy Target" object, which can readily be placed throughout the arena at strategic points. Whenever the subject touches a

Dummy Target, the time and location of the touch is recorded, allowing a complete reconstruction of their search path and timing from the data automatically collected by fAARS-Lite.

The Geometric Arena

Inspired by Cheng (1986), the basic geometric arena paradigm pairs specific feature cues (colored wall panels) with enclosed arenas of specific shapes. Subjects are required to find hidden reinforcers, typically in the corners of these environments, by learning which combination of geometric and feature cues correspond to reinforcement. The experimental manipulation changes the relative configuration of these cues, allowing experimenters to place the cue types in conflict or remove one type of cue altogether. Unlike the previous two tasks, these studies frequently involve multiple arenas being tested simultaneously (for instance, subjects trained in a rectangle with a feature in location X might be simultaneously tested in a rectangle with a feature in location Y, a rectangle without features, or a square arena with a feature in location X). Studies of this sort have been carried out with rats, chicks, pigeons, fish, and both human and non-human primates (reviewed in Cheng, 2008 and Cheng & Newcombe, 2005).

Bringing this task to life in fAARS-Lite is very much like combining the previous two tasks. Within the client's virtual world, each training condition consists of its own enclosed arenas, composed of four walls in a particular configuration and a unique Destination in the center. A unique Target is placed in each corner, along with any visible feature cues appropriate to the experiment. As in the experiments above, the testing condition consists of a copy of the training environments with the appropriate cues moved or removed.

The server-side implementation is identical to the radial arm maze with one exception: During the testing phase, multiple arenas are specified as valid. The fAARS-Lite engine automatically generates a randomized presentation schedule for these arenas, such that each arena is presented a specified number of times but randomly interspersed with visits to other, simultaneous arenas (in order to prevent order effects). As in the above examples, "resetting" - travel between a chosen Target and the next Destination - is masked from the subject's view via the Signal's feedback screen.

fAARS-Lite has been used successfully in the literature to model geometric arena tasks. Lubyk, Dupuis, Gutiérrez and Spetch (2012) used fAARS-Lite to implement a human analogue of a geometric arena task originally designed for chicks (Tommasi & Polli, 2004) and pigeons (Lubyk & Spetch, 2012). In this paradigm, subjects were trained to locate specific corners in parallelogram-shaped arenas. These corners are identified through unique wall-length configurations (i.e. long left wall, short right wall) as well as their angular amplitude (i.e. acute or obtuse). Subjects were then tested in rectangular arenas (in which all four corners are the same angle), rhomboid arenas (in which the wall lengths are all identical), and mirror-image parallelogram arenas (in which the angle and wall-length cues from training are in conflict). fAARS-Lite was configured as described above, with Targets in each arena corner, Destinations set in each arena center, and Signals configured to provide “correct/incorrect” feedback messages during training, and “no feedback” messages during testing. Up to six subjects were tested simultaneously, with 99 human subjects in all. Lubyk et al. (2012) found that subjects had encoded both wall-length and angular information, and that the latter was weighted more heavily in conflict tests. These results with humans are comparable to existing animal data with both species of bird discussed above. A direct comparison between all three species is straightforward, as these experiments are direct analogues of each other, and because non-immersive virtual methods produce results comparable to immersive methods (Kelly & Gibson, 2007; Sturz et al., 2009).

Conclusion

The fAARS-Lite platform provides an effective tool for comparative cognition researchers interested in human performance on spatial tasks. Constructed completely from free and well-documented open-source software, and organized through a graphical user interface (instead of being programmed using text-based commands), it is easy to set up and deploy without significant investment in funding, time, or expertise. The platform’s logic is very basic, and can be used without modification to recreate any task that involves sequential visitation. This article provided example arrangements for three classic spatial paradigms – the radial arm maze (Olton &

Samuelson, 1976), the Morris water maze (Morris, 1984), and the geometric arena (Cheng, 1986) – but a potentially limitless variety of tasks can be described using the same logic.

An even wider range of tasks can be modeled with slight modification of some of the component systems. For example, Target objects currently activate if a participant moves into contact with them, representing movement into a particular region of space. While this is sufficient for many tasks, others require more interaction with the environment. This could be modeled in fAARS-Lite by adjusting individual Target scripts to activate in response to a different action, such as a mouse click.

References

Abrahams, S., Pickering, A., Polkey, C. E., & Morris, R. G. (1997). Spatial memory deficits in patients with unilateral damage to the right hippocampal formation. *Neuropsychologia*, *35*(1), 11–24.

Brown, M. F., & Demas, G. E. (1994). Evidence for spatial working memory in honeybees (Apis mellifera). *Journal of Comparative Psychology*, *108*(4), 344–352.

Cheng, K. (1986). A purely geometric module in the rat’s spatial representation. *Cognition*, *23*(2), 149–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3742991

Cheng, K. (2008). Whither geometry? Troubles of the geometric module. *Trends in cognitive sciences*, *12*(9), 355–61. doi:10.1016/j.tics.2008.06.004

Cheng, K., & Newcombe, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic bulletin & review*, *12*(1), 1–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15945200

Cole, M. R., & Chappell-Stephenson, R. (2003). Exploring the limits of spatial memory in rats, using very large mazes. *Learning & behavior*, *31*(4), 349–68. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14733483

DiGian, K. a, & Zentall, T. R. (2007). Pigeons may not use dual coding in the radial maze analog task. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(3), 262–72. doi:10.1037/0097-7403.33.3.262

Gutiérrez, L. A. (2012). *The fAARS Platform For Augmented Alternate Reality Services and Games*. University of Alberta. Retrieved from https://era.library.ualberta.ca/public/view/item/uuid:b8d36cef-585d-4010-8124-9a22c1384087

Hamilton, D. a, Driscoll, I., & Sutherland, R. J. (2002). Human place learning in a virtual Morris water task: some important constraints on the flexibility of place navigation. *Behavioural brain research*, *129*(1-2), 159–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11809507

Kelly, D. M., & Bischof, W. F. (2005). Reorienting in images of a three-dimensional environment. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(6), 1391–403. doi:10.1037/0096-1523.31.6.1391

Kelly, D. M., & Gibson, B. M. (2007). Spatial Navigation: Spatial Learning in Real and Virtual Environments. *Comparative Cognition and Behaviour Reviews*, *2*, 111–124. doi:10.3819/ccbr.2008.20007

Lubyk, D. M., Dupuis, B., Gutiérrez, L. A., & Spetch, M. L. (2012). Geometric orientation by humans: angles weigh in. *Psychonomic Bulletin & Review*, *19*(3), 436–442. doi:10.3758/s13423-012-0232-z

Lubyk, D. M., & Spetch, M. L. (2012). Finding the best angle: pigeons (Columba livia) weight angular information more heavily than relative wall length in an open-field geometry task. *Animal cognition*, *15*(3), 305–12. doi:10.1007/s10071-011-0454-x

MacDonald, S. (1994). Gorillas'(Gorilla gorilla gorilla) spatial memory in a foraging task. *Journal of Comparative Psychology*, *108*(2), 107–113.

Morris, R. (1984). Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of neuroscience methods*, *11*(1), 47–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6471907

Mou, W., Biocca, F., Owen, C., Tang, A., & Lim, L. (2004). Frames of reference in mobile augmented reality displays. *Journal of Experimental Psychology: Applied*, *10*(4), 238–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15598121

Newcombe, N. S., Ratliff, K. R., Shallcross, W. L., & Twyman, A. D. (2010). Young children's use of features to reorient is more than just associative: further evidence against a modular view of spatial processing. *Developmental science*, *13*(1), 213–20. doi:10.1111/j.1467-7687.2009.00877.x

Newman, M. C., & Kaszniak, A. W. (2010). Spatial Memory and Aging : Performance on a Human Analog of the Morris Water Maze. *Aging, Neurophysiology, and Cognition*, *7*(2), 86–93.

Olton, D., & Samuelson, R. J. (1976). Remembrance of Places Passed : Spatial Memory in Rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(2), 97–116.

O'Connor, R. C., & Glassman, R. B. . (1993). Human performance with a seventeen-arm radial maze analog. *Brain Research Bulletin*, *30*(1-2), 189–191. doi:10.1016/0361-9230(93)90058-J

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. London: Clarendon.

Rahman, Q., Abrahams, S., & Jussab, F. (2005). Sex differences in a human analogue of the Radial Arm Maze: the "17-Box Maze Test". *Brain and cognition*, *58*(3), 312–7. doi:10.1016/j.bandc.2005.03.001

Shettleworth, S. J. (2010). *Cognition, evolution, and behaviour*. Oxford, UK: Oxford University Press.

Skelton, R. W., Ross, S. P., Nerad, L., & Livingstone, S. a. (2006). Human spatial navigation deficits after traumatic brain injury shown in the arena maze, a virtual Morris water maze. *Brain injury*, *20*(2), 189–203. doi:10.1080/02699050500456410

Spetch, M. L., & Edwards, C. a. (1986). Spatial memory in pigeons (Columba livia) in an open-field feeding environment. *Journal of Comparative Psychology*, *100*(3), 266–278. doi:10.1037//0735-7036.100.3.266

Sturz, B. R., & Bodily, K. D. (2010). Encoding of variability of landmark-based spatial information. *Psychological research*, *74*(6), 560–7. doi:10.1007/s00426-010-0277-4

Sturz, B. R., Bodily, K. D., Katz, J. S., & Kelly, D. M. (2009). Evidence against integration of spatial maps in humans: generality across real and virtual environments. *Animal cognition*, *12*(2), 237–47. doi:10.1007/s10071-008-0182-z

Talbot, K. J., Legge, E. L. G., Bulitko, V., & Spetch, M. L. (2009). Hiding and searching strategies of adult humans in a virtual and a real-space room. *Learning and Motivation*, *40*(2), 221–233. doi:10.1016/j.lmot.2009.01.003

Tolman, E. C. (1948). Cognitive maps in rats & in men. *Psychological Review*, *55*(4), 189–208.

Tommasi, L., & Polli, C. (2004). Representation of two geometric features of the environment in the domestic chick (Gallus gallus). *Animal cognition*, *7*(1), 53–9. doi:10.1007/s10071-003-0182-y

Wilkie, W. M., & Slobin, P. (1983). Gerbils in Space: performance on the 17-arm maze. *Journal of the Experimental Analysis of Behavior*, *40*, 301–312.

Zhao, M., Zhou, G., Mou, W., Hayward, W. G., & Owen, C. B. (2007). Spatial updating during locomotion does not eliminate viewpoint-dependent visual object processing. *Visual Cognition*, *15*(4), 402–419. doi:10.1080/13506280600783658

Chapter 4

Get Out Of The Corner: The Effect Of Location Type And Number On Perceptron And Human Reorientation

Brian Dupuis & Michael R.W. Dawson

Department of Psychology, University of Alberta

Author Note:

[This chapter is a modified form of a manuscript being prepared for submission by Dupuis and Dawson.]

Introduction

Any mobile agent, capable of navigating through its world, must be able to find its bearing and orientation to do so. Researchers have developed a number of paradigms to investigate this ability, the foremost of which is the ‘reorientation task’ (Cheng, 1986). Within a reorientation task, agents are placed inside a controlled arena which contains a specific set of cues, and are trained to search a particular location for reinforcement. Following testing, the arena is reconfigured, and the changes in agent search patterns are recorded.

A typical reorientation task uses a quadrilateral arena, with four locations of interest (one in each corner). While some exceptions exist (Cheng, 1986; Newcombe, Ratliff, Shallcross, & Twyman, 2010), the overwhelming majority of reorientation experiments conform to this structure (Cheng & Newcombe, 2005). The emphasis on corners has some desirable properties – for instance, in a rectangle, corners that are diagonally opposite are geometrically identical. Observe, in Figure 4-1A, Locations 1 and 3 both have identical geometry (long wall on the right, short wall on the left, both walls joined at a 90° angle). Because of this, if an agent was choosing a location on the basis of geometry alone, it would have the same likelihood of visiting Location 1 as it would of visiting Location 3. This led to the discovery that rats processed geometric cues even when they could be completely ignored (Cheng, 1986; Gallistel, 1990). In a typical study (for instance, Wall, Botly, Black, & Shettleworth, 2004), Location 1 of Figure 4-1A would be reinforced when visited, to indicate that it was the ‘correct’ location. Furthermore, a unique landmark – a non-geometric identifier – would be placed at that location. This landmark provides sufficient information for an animal to learn the ‘correct’ location; geometric information is not required, and is in fact less reliable, because the geometric features at Location 1 are also present at the non-reinforced Location 3. Nevertheless, if, after training, the rat is placed in a new arena in which the ‘correct’ landmark has been moved to Location 2, its behaviour will typically indicate that geometric cues were encoded. That is, it will have a high likelihood of visiting Location 2 in the new arena (indicating that non-geometric cues were learned), but will also visit Locations 1 and 3 (indicating that geometric cues were also learned). Such results led to the development of the geometric module theory (Cheng, 1986; Gallistel, 1990), which is based on the assumption that geometric cues are processed independently of non-geometric cues, and that the processing of geometric cues is mandatory.

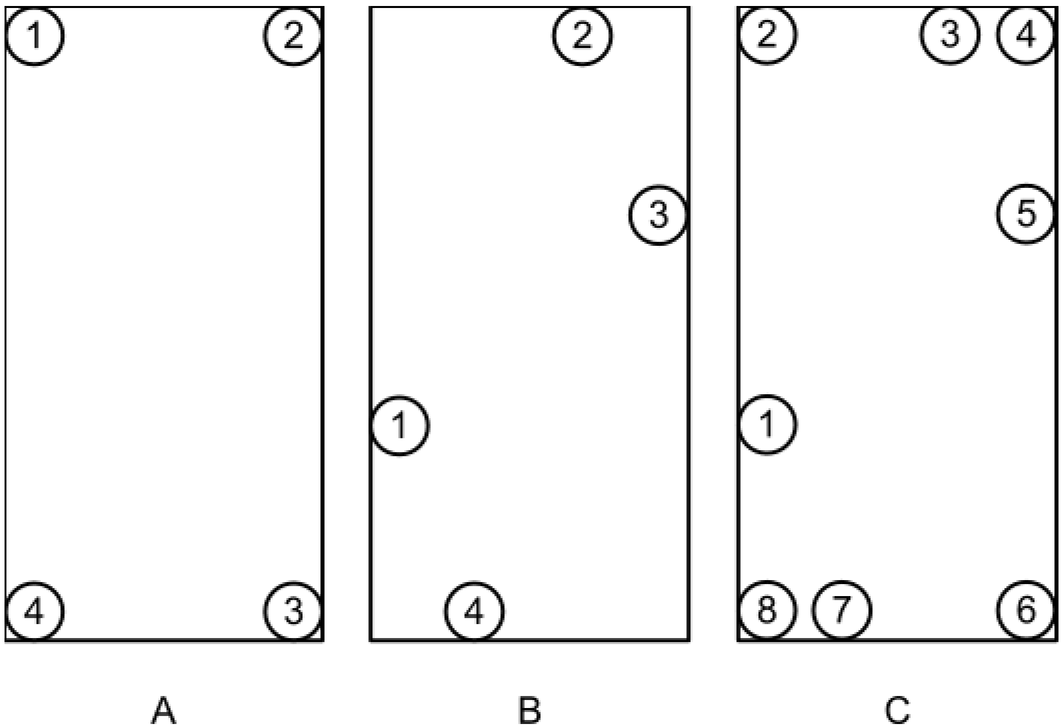

Figure 4-1: Possible configurations of targets in a rectangular arena. Many tasks concern themselves solely with corners (A), but similar relationships (i.e. long wall to the left of the target, short wall to the right of the target) can occur in other locations, such as along walls (B). Such arenas therefore need not be limited to four locations of interest (C).

The geometric module theory is an example of an insight into navigation provided by the reorientation task. This fairly straightforward task has been a fruitful source of information about navigation, and has been used to study a wide variety of organisms, including ants, fish, rats, birds, and humans (review in Cheng, 2008; Cheng & Newcombe, 2005). While this informative task is straightforward to describe and provides data that is easily analyzed, it is important to realize that it is expensive to conduct. Particularly when using animal subjects, an experiment requires considerable commitment of resources, because subjects must be run individually and it takes a fair amount of training for a subject to learn the ‘correct’ location before being placed in a novel arena. It would be convenient if there was a less expensive medium in which to explore the reorientation task, with the aim of discovering interesting hypotheses that could then be tested with a traditional (and more expensive) experiment.

Computer simulations are one less-expensive medium that can be used to explore domains of interest. Lewandowsky (1993) has pointed out that computer simulations can provide several advantages for theory development in cognitive science. These include formalizing a theory in such a way that rigor is improved, and providing more precise tools for studying concepts of interest. Additionally, a working computer simulation can be used to reveal tacit assumptions hidden within a theory. Finally, a computer simulation can itself lead to serendipitous findings, particularly when it is presented novel situations. It would seem that if one had a plausible computer simulation for the reorientation task, then it could be used to explore new situations with ease, and could possibly generate unexpected predictions. These new simulation-based predictions could then be tested with a traditional experiment, particularly if a researcher felt that the predictions were interesting enough to warrant the experiment's expense.

Fortunately, a plausible computer simulation for reorientation has been proposed (Dawson, Kelly, Spetch, & Dupuis, 2010) in the form of a simple artificial neural network called a perceptron (Rosenblatt, 1962). In its standard form, a perceptron consists of a single bank of input units which numerically encode patterns of stimuli; these input units are linked via weighted connections to an output unit, which transforms this weighted net input signal into response behaviour. The strength of the connection weights is then updated following a specified learning rule, designed to minimize the difference between the output unit's activity and the desired response to that particular input pattern. A perceptron trained with a standard learning algorithm has been shown to generate most of the interesting regularities found in reorientation task behaviour (Dawson, Kelly, et al., 2010). An operant perceptron model of reorientation, which uses a more psychologically plausible learning algorithm, has also been proposed, and has shown some promising results (see Chapter 1 and Dawson, Dupuis, Spetch, & Kelly, 2009). The purpose of the current paper is to illustrate how an operant perceptron can be used to explore reorientation by observing the model's behaviour when novel reorientation paradigms are simulated. We demonstrate that this kind of computer simulation can generate interesting predictions which can then be tested using more traditional experimental methodologies.

Part of the power of computer simulations is that, by revealing tacit assumptions about the processes responsible for phenomena of interest (Lewandowsky, 1993), they also provide the means to challenge, or even negate, those assumptions. In the case of reorientation, while the geometric module theory had a strong early impact for many years, it has recently been questioned, with some researchers arguing for it to be abandoned completely (Cheng, 2008; Twyman & Newcombe, 2010). If reorientation is accomplished without the use of a geometric module, then what mechanisms might instead be responsible? One alternative to the geometric module is an appeal to general principles of associationist learning (Miller & Shettleworth, 2007). According to this view, there is no geometric module, but both geometric and non-geometric features are treated in the same manner as being sources of information (i.e. cues). Agents use standard learning procedures to associate the various available cues (both geometric and non-geometric) at a location with the likelihood of being rewarded at that location. A model employing this approach has been shown to be capable of simulating many reorientation task regularities without an appeal to the geometric module (Miller, 2009; Miller & Shettleworth, 2007, 2008). However, serious empirical and theoretical problems with this model have been identified (see Chapter 1 and Dawson, Kelly, Spetch, & Dupuis, 2008). The standard perceptron corrects these problems, and also models reorientation regularities (Dawson, Kelly, et al., 2010). The operant perceptron captures the same regularities as the standard perceptron, but does so with an even more realistic conception of learning in the reorientation paradigm (Chapter 1). Despite their mathematical differences, as reactions to geometric modularity both the purely associationist and perceptron models were developed in the spirit of challenging the assumptions about the processes underlying reorientation.

Exploring Assumptions Used To Define The Reorientation Task

Because perceptrons are plausible computer simulations for studying reorientation, they enable the exploration of other challenges, concerning not just assumptions about reorientation processes, but also assumptions about the reorientation paradigm itself. For example, although the prototypical reorientation paradigm (Cheng, 1986) employed a grid spanning the entire arena for measurement, it was noted earlier that a common feature of most modern reorientation experiments is the emphasis on corners as the only locations of interest. Might performance on the reorientation task be affected if agents were trained to go to locations that are not at corners? Are corner locations special in some way?

According to associationist theories of reorientation, corner locations should not be special. These theories, based on the work of Rescorla and Wagner (1972), posit a view of learning largely based on cue competition. Applied in to the reorientation task, this suggests that a location is merely a collection of cues that can be exploited as signals of potential reinforcement. From this perspective, there should be no fundamental behavioural difference between a location of interest at a 90° corner, and one along a 180° wall. In effect, a location under consideration along a single wall "divides" the wall into a left and right segment, exactly as a corner does, with only the angle of intersection distinguishing them. Under this associative viewpoint, it would appear that locations of interest need not be constrained to corners.

Another tacit assumption that guides experimental studies of reorientation concerns the number of target locations. The vast majority of reorientation studies have used quadrilateral arenas, typically rectangles or squares (review in Cheng, 2008; Cheng & Newcombe, 2005), and less commonly in kites (Dawson, Kelly, et al., 2010; Pearce, Good, Jones, & McGregor, 2004) or parallelograms (Lubyk & Spetch, 2012; Tommasi & Polli, 2004). In these studies, the corners of these arenas have been used as target locations, and therefore all of these experiments have studied reorientation using four different locations. Furthermore, this paradigm only makes available two different instances of geometric cues (long wall on the left and short wall on the right of a corner, short wall on the left and long wall on the right of a corner). Only a handful of studies have used arenas that are not quadrilateral (i.e. Newcombe et al., 2010), and have therefore made available more than 4 target locations (again assuming that locations of interest are always positioned at arena corners). To our knowledge, no experimental studies of reorientation have explicitly compared situations in which the number of target locations has been systematically varied.

However, it is important to study the effect of varying the number of target locations, because some theories predict that this variable should affect learning in the reorientation task. For example, Miller and Shettleworth's (2007, 2008) associative model uses a measure of the probability of an agent visiting a particular location in its learning equations. This measure is expressed as the net attractiveness of the location as a proportion of the total net attractiveness of every location. This proportion will obviously be affected by the number of locations that are summed in its denominator. All else being equal, this will predict (at least initially) one-half the normal rate of learning for tasks with 8 locations relative to tasks with 4 locations[5].That is, the Miller-Shettleworth model predicts that learning the reorientation task will slow down as the number of possible locations of interest increases. Do other models, such as the perceptron, also make this prediction?

Furthermore, as more locations are added to a standard reorientation arena (Figure 4-1C) a greater variety of geometric cues must be processed. For instance, there are four different geometric configurations that can be distinguished in Figure 4-1C, which again would be expected to slow reorientation learning in a theory like Miller and Shettleworth's (2007, 2008) associative model. Does the number of target locations affect human learning in the reorientation paradigm, and does it do so in a fashion predicted by computer simulations? The current paper represents an attempt to begin the exploration of such questions.

The purpose of the current manuscript is to investigate the two main issues raised above. First, we attempt to evaluate the role that the nature of a location – at a corner or at a wall – has on reorientation behaviour. Second, we investigate the impact that changing the number of locations of interest has on reorientation behaviour. These distinctions are depicted in Figure 4-1. Figure 4-1A illustrates the typical position of possible target locations in a standard reorientation task that uses a rectangular arena. Figure 4-1B provides an analogous arena, but one in which the locations of interest are not found at corners. Figure 4-1C shows how one can combine the first two arenas into a third that has 8 locations of interest instead of the typical 4.

[5] Due to other serious mathematical flaws with the Miller and Shettleworth (2007, 2008), detailed at length in Chapter 1, we will not go into further detail with this model here.

This paper proceeds as follows: It begins by using an operant perceptron (see Chapter 1 and Dawson, Dupuis, Spetch, & Kelly, 2009) to simulate the reorientation task in the various arenas illustrated in Figure 4-1. These simulations are used to make predictions about the effects of type of location and of number of locations on reorientation behaviour. The results of these simulations provide two key predictions: 1) corner locations are not inherently special, and 2) doubling the number of target locations has a negligible effect on the speed at which the model learns to reorient. Next, we report the results of testing these predictions using human subjects in a virtual world (see Chapter 3). Finally, we explore the similarities and differences between the associationist model and the human data. We argue that the operant perceptron is a useful source of predictions that can be supported by experimental data. As a result, the operant perceptron appears to provide a medium in which reorientation can be plausibly explored for the purpose of seeking surprising and interesting results that can later become the focus of traditional experimentation.

Simulation

From the perspective of theories of reorientation that appeal to a geometric module (Gallistel, 1990), angle information present at a corner is typically viewed as a global, geometric property. However, from an associationist perspective, a "corner" could be perceived as a visually salient "focal point" that serves as a reference, with the angle simply being a (local) feature of that location. The intersection between walls provides a distinct boundary, from which the length of a wall can be measured. For instance, in Figure 4-1B, Location 1 sits at the junction of a short wall on its left and a long wall on its right, with an intersection angle of 180°.

With this in mind, we devised a method of representing any location along the edge of an arena that treats angle information as a feature. This representation permitted us to present locations to perceptrons, even when these locations were not at a corner in a reorientation arena.

The Perceptron

Perceptron Reorientation

As was noted earlier, a perceptron (Rosenblatt, 1962) is a simple artificial neural network in which a set of input units are directly connected to an output unit via weighted connections. The input units represent stimuli; their activation causes signals to be sent through the weighted connections to produce a response in the perceptron's output unit. Feedback can be provided to the network about its response so that it can modify its connection weights. This permits the perceptron to learn to generate a desired response to each stimulus in a set of training patterns.

To simulate the reorientation task, each location of interest in an arena is represented as a stimulus in the set of training patterns. For each of these patterns, input unit activity is used to represent which cues (geometric and non-geometric) are present at a particular location. If a location is deemed to be 'correct', then the perceptron is reinforced when that location's cues are presented. If a location is not deemed to be 'correct', then the perceptron is not reinforced when that location's cues are presented. In other words, the perceptron is trained to produce an activity of 1 to sets of cues corresponding to 'correct' locations, and an activity of 0 to sets of cues corresponding to 'incorrect' locations.

In order to train the perceptron to learn to reorient in a particular arena, one must make design decisions about how to represent the available cues, and about the learning rule that is used to modify the network's connection weights. The details of these design decisions are provided below.

Defining the task: Stimuli

Each location identified in Figure 4-1 can be defined as a collection of properties, which are presented to the perceptron as a pattern of unary-coded inputs. That is, each of the perceptron's input units encodes the presence or absence of a specific cue. Each of these units is turned on (activated with a value of 1) when the property it encodes is present, and is turned off (activated with a value of 0) when that property is absent. In the current simulation, each location of interest is defined by three types of cues: the length of the walls on either side of the location, the angle between the walls where they join, and the kind of local landmark that can be present at the location. 17 different input units were used to represent the possible values of these cues as is summarized in Table 4-1.

Table 4-1

Encoding of a location's properties and the agent's response using an operant perceptron.

Unit	Codes For	Encoding	Example Values	Example Encoding
Inputs 1-2	Angle at target	Unary coding, 2 units	90°	1 0
			180°	0 1
Inputs 3-11	Color of target	Unary coding, 9 units	Red	1 0 0 0 0 0 0 0 0
			Green	0 1 0 0 0 0 0 0 0
			Yellow	0 0 1 0 0 0 0 0 0
			Blue	0 0 0 1 0 0 0 0 0
			Orange	0 0 0 0 1 0 0 0 0
			Black	0 0 0 0 0 1 0 0 0
			Brown	0 0 0 0 0 0 1 0 0
			Purple	0 0 0 0 0 0 0 1 0
			White	0 0 0 0 0 0 0 0 1
Inputs 12-17	Configuration of wall lengths at target	Unary coding, 6 units	Left 3, Right 6	1 0 0 0 0 0
			Left 6, Right 3	0 1 0 0 0 0
			Left 2, Right 1	0 0 1 0 0 0
			Left 2, Right 4	0 0 0 1 0 0
			Left 4, Right 2	0 0 0 0 1 0
			Left 1, Right 2	0 0 0 0 0 1

The angle units (1-2) identify the angle of intersection at the location. These units have one value for locations at corners (90° angle), and another value for locations along walls (180° angle). The feature units (3-11) represent the collection of non-geometric properties present at a given location. For parsimony with the experiments described later in the manuscript, these units are named after colors; as such, these units can be thought of as representing the color of an object at the location.

The length configuration units (12-17) represent the specific set of wall-length properties present at a location. For example, one unit is turned on for a location at the intersection of a wall of length 3 with a wall of length 6, while another might be turned on if the location lies between walls of length 2 and 1. This is an extension of Miller and Shettleworth's (2007, 2008) representation for specific geometries that allows for a number of possible configurations – up to six in the current simulation. This is required when more than 4 locations of interest are used (Figure 4-1C).

This particular set of design decisions defines this encoding as a purely local code – that is, each pattern contains only information present at the location it represents, and no information from any other location. Similarly, this encoding contains no global representation of the arena, either explicitly (i.e. a principal axis, Cheng & Gallistel, 2005) or implicitly (as in Miller and Shettleworth's (2007, 2008) model summing across all locations), save for the number of patterns presented.

Defining the task: Response

This simulation includes a single output unit that uses the logistic activation function (Dawson, 2008) to convert the total weighted signal coming from the input units into a response that can range between 0 and 1. For locations that are reinforced, the perceptron is trained to turn on (output activity = 1); for locations that are not reinforced, the perceptron is trained to turn off (output activity = 0). Because, during learning, perceptron activity falls in the continuous range between 0 and 1, at any moment in time, the perceptron's output can be interpreted as its estimation of the likelihood of receiving reinforcement at the current location (see Chapter 1, also Dawson, Dupuis, Spetch, & Kelly, 2009).

Training Method

A perceptron's response to particular patterns of stimuli is not perfect; each generated response differs from a desired response by some error amount. This error is then used by a learning rule to adjust the perceptron's connection weights such that subsequent presentations of that pattern of stimuli produce a smaller error. Here, we employ the gradient-descent learning rule (Dawson, 2004, 2008), which has desirable properties when working a logistic perceptron response.

This output response provides a critical distinction between neural network models and traditional associative models in the style of Rescorla and Wagner (1972). The perceptron's output activity allows it to convert *associative strength* of assorted cues into a model of *behaviour*. This stage is absent from traditional associative models. Not only does this difference allow perceptrons to produce different predictions from formally-equivalent associative models (Dawson, 2008), it also allows us to adjust the model's learning to reflect different patterns of behaviour.

In the current model, this measure of behaviour is used to adjust the perceptron's learning from classical conditioning to operant conditioning, where it is allowed to 'choose' whether or not to investigate a particular location, and this investigation (rather than rote presentation) governs its learning. Instead of updating connection weights after every pattern of cues is presented, the perceptron's output response to that pattern is used as the *probability* of updating weights on this presentation. At each presentation, a random number between 0 and 1 is generated and compared to the output response; if the random number exceeds the output response, the connection weights are not updated, and the next pattern is presented. This algorithm is detailed at length in Chapter 1.

Simulation Specification

Training

The current simulation includes two experimental conditions: one with four locations of interest, and one with eight locations of interest. Each location is present at either a wall or a corner, and contains a unique feature cue (i.e. a colored object). Within each condition, networks are trained to investigate just one location; this location is reinforced, while all others are not reinforced. The reinforced location could be present at a wall or a corner, producing a 2 (four-vs.-eight) x 2 (corner-vs.-wall) design.

All networks were initialized with all biases and connection weights equal to zero, and were trained with a learning rate of 0.1. Five networks in each condition were trained to convergence. For counterbalancing, two possible reinforcement locations were used in each simulation; for example, in the four-location, corner-goal task (Figure 4-1A), one group of networks is reinforced at Location 1, while another is reinforced at Location 2. No appreciable difference was found between these groups, so their results are reported together here. (That is, each value is averaged from ten networks.)

Testing

Testing the perceptron involves presenting patterns of cues corresponding to transformed arenas, and measuring the perceptron's output response to these novel patterns. Due to the operant nature of its training algorithm, the perceptron's output response is both its estimation of reward likelihood at the location given the cues presented, and its likelihood of visiting that location.

There are two types of transformed arenas common to reorientation studies: affine transformations and 'featureless' transformations. Affine transformations place feature cues and geometry cues in conflict with each other: a chosen location could be consistent with the geometry present during training, or the features present during training. For instance, in Figure 4-1A, a subject might find reinforcement at Location 1, along with a unique feature. When placed in an arena with an affine transformation, that unique feature might now be present at Location 2. Location 2 is consistent with training in terms of features, while Locations 1 and 3 are consistent with training in terms of geometry. Meanwhile, a featureless transformation replaces all unique feature cues with indistinguishable ones, forcing the model to base its decisions solely on encoded geometry. In the current simulation, we simply turn off all feature units which were present during training, and activate a novel "white" feature unit in their place.

With 4 locations, we can also test for generalization across angle cues by observing a corner-trained network's response to a wall-locations-only arena (that is, a network trained in Figure 4-1A but tested in Figure 1B), and vice versa. In this scenario, each location now appears with novel angle and length configuration cues, as opposed to an affine transformation which has novel length configurations but consistent angles. As these two conditions do not share exact wall lengths, no choice can be consistent with wall length geometry from training.

With 8 locations, one can also do a "partial" transformation. While affine-transformed arenas have consistent angle information (targets that were present at corners are present at corners during testing), a partial transform places them in conflict. Both transformations have novel wall length geometries compared to the training condition.

Following training, each network was presented with probe trials in three transformed arenas: for 4 locations, these were affine, generalized, and featureless arenas. For 8 locations, these were affine, partial, and featureless arenas. Each network's output responses were recorded for each of these locations; these responses were averaged across the five networks present at each condition.

In reorientation task literature, it is common to report responses in terms of the frequency with which each location is chosen. However, the perceptron responds to each location individually, producing the probability of choosing to act at that specific location; these probabilities need not sum to 1 across all locations within an arena. In order to convert the former into the latter, we divided the response to a specific location by the sum of responses to all locations within a given arena; this method has previously been used to successfully predict several key reorientation behaviour regularities (Dawson, Kelly, et al., 2010).

Results

Across all conditions, networks converged after an average of 4810 presentations of the training set (a single presentation of each pattern in the training set in a random order is called a 'sweep'), with the fastest training occurring after 4614 sweeps of training and the slowest training requiring 4973 sweeps.

Network Responses

The network model's responses to each location in each transformed arena, expressed both as response activity and as choice frequencies, are reported in Tables 4-2 and 4-3. The tables include a summary of human responses in similar conditions in experiments that were inspired by the simulation results. The human responses in the table are covered in more detail when the human experiments are discussed, below.

Table 4-2

Average responses of operant perceptrons and human subjects to the 4-location simulation.

Wall configuration	Network	Human

Task	Type	Transform	Loc	Color	Angle	Left	Right	Activity	Frequency	Frequency
Four	Corner	Affine	1	Blue	90*	6*	3*	0.36	0.24	0.06
			2	Green*	90*	3	6	0.75	0.50	0.89
			3	Red	90*	6*	3*	0.36	0.24	0.03
			4	Yellow	90*	3	6	0.01	0.01	0.02
		Generalized	1	Blue	180	2	4	0.26	0.17	0.04
			2	Green*	180	2	1	0.94	0.62	0.87
			3	Red	180	2	4	0.26	0.17	0.04
			4	Yellow	180	2	1	0.06	0.04	0.04
		Featureless	1	White	90*	6*	3*	0.52	0.42	0.26
			2	White	90*	3	6	0.11	0.08	0.19
			3	White	90*	6*	3*	0.52	0.42	0.30
			4	White	90*	3	6	0.11	0.08	0.25
	Wall	Affine	1	Yellow	180*	2*	4*	0.37	0.25	0.16
			2	Blue*	180*	2	1	0.75	0.50	0.68
			3	Green	180*	2*	4*	0.36	0.24	0.11
			4	Red	180*	2	1	0.01	0.01	0.05
		Generalized	1	Green	90	6	3	0.26	0.22	0.02
			2	Red	90	3	6	0.06	0.05	0.05
			3	Yellow	90	6	3	0.26	0.22	0.03
			4	Blue*	90	3	6	0.60	0.51	0.49
		Featureless	1	White	180*	2*	4*	0.53	0.42	0.26
			2	White	180*	2	1	0.11	0.08	0.22
			3	White	180*	2*	4*	0.53	0.42	0.35
			4	White	180*	2	1	0.11	0.08	0.29

* Cue type reinforced during training

Table 4-3

Average responses of operant perceptrons and human subjects to the 8-location simulation.

						Wall configuration		Network		Human
Task	Type	Transform	Loc.	Color	Angle	Left	Right	Activity	Frequency	Frequency
Eight	Corner	Affine	1	Black	180	2	4	0.04	0.03	0.01
			2	Blue	90*	4*	2*	0.36	0.22	0.16

		3	Purple	180	2	1	0.04	0.03	0.02
		4	Green*	90*	1	2	0.74	0.45	0.58
		5	Orange	180	2	4	0.04	0.03	0.00
		6	Red	90*	4*	2*	0.36	0.22	0.18
		7	Brown	180	2	1	0.04	0.03	0.04
		8	Yellow	90*	1	2	0.01	0.01	0.01
	Partial	1	Blue	180	2	4	0.03	0.02	0.01
		2	Purple	90*	4*	2*	0.42	0.25	0.19
		3	Green*	180	2	1	0.61	0.36	0.65
		4	Orange	90*	1	2	0.07	0.04	0.00
		5	Red	180	2	4	0.03	0.02	0.00
		6	Brown	90*	4*	2*	0.42	0.25	0.11
		7	Yellow	180	2	1	0.01	0.00	0.03
		8	Black	90*	1	2	0.07	0.04	0.01
	Featureless	1	White	180	2	4	0.05	0.04	0.00
		2	White	90*	4*	2*	0.49	0.35	0.26
		3	White	180	2	1	0.05	0.04	0.06
		4	White	90*	1	2	0.09	0.07	0.15
		5	White	180	2	4	0.05	0.04	0.01
		6	White	90*	4*	2*	0.49	0.35	0.26
		7	White	180	2	1	0.05	0.04	0.09
		8	White	90*	1	2	0.09	0.07	0.17
Wall	Affine	1	Black	180*	2*	4*	0.36	0.22	0.11
		2	Blue	90	4	2	0.04	0.03	0.00
		3	Purple*	180*	2	1	0.74	0.45	0.76
		4	Green	90	1	2	0.04	0.03	0.00
		5	Orange	180*	2*	4*	0.36	0.22	0.13
		6	Red	90	4	2	0.04	0.03	0.00
		7	Brown	180*	2	1	0.01	0.01	0.00
		8	Yellow	90	1	2	0.04	0.03	0.00
	Partial	1	Blue	180*	2*	4*	0.42	0.25	0.18
		2	Purple*	90	4	2	0.61	0.36	0.71
		3	Green	180*	2	1	0.07	0.04	0.00
		4	Orange	90	1	2	0.03	0.02	0.00
		5	Red	180*	2*	4*	0.43	0.25	0.11

	6	Brown	90	4	2	0.01	0.00	0.00
	7	Yellow	180*	2	1	0.07	0.04	0.00
	8	Black	90	1	2	0.03	0.02	0.00
Featureless	1	White	180*	2*	4*	0.50	0.35	0.29
	2	White	90	4	2	0.05	0.04	0.15
	3	White	180*	2	1	0.09	0.07	0.13
	4	White	90	1	2	0.05	0.04	0.00
	5	White	180*	2*	4*	0.50	0.35	0.18
	6	White	90	4	2	0.05	0.04	0.14
	7	White	180*	2	1	0.09	0.07	0.07
	8	White	90	1	2	0.05	0.04	0.03

* Cue type reinforced during training

The first major prediction generated by the model is that there does not appear to be a significant difference in reorientation behaviour between subjects trained with locations in corners and subjects trained with locations along walls. In conditions with 4 locations of interest, and 8 locations of interest, the perceptron converged after a similar number of sweeps of training. Furthermore, in all cases, the same broad pattern of behaviour holds: the perceptron responds most strongly to locations containing the (unique) feature cue present during training, but that cue did not block the encoding of either geometric cue. That is, even within the Featureless arena, the perceptron still estimates that locations with the same wall length configuration and/or angle amplitude as the training location have a greater likelihood of reward than locations missing those cues.

Additionally, the perceptron produces characteristic “rotational error” behaviour common to reorientation tasks (Cheng, 1986) – that is, where features and geometry conflict in the same arena, the perceptron responds to the feature more frequently than to any other single location, but taken as a whole, locations with correct geometry are chosen with higher frequency. This pattern appears in both the 4-location and 8-location tasks. Furthermore, it occurs even if the angle information changes between conditions – for instance, the Generalized arenas in the 4-location task still produce this pattern, even though the exact configuration of geometries present in this condition are novel.

Connection Weights

To understand why these networks behave in this manner, we turn next to their connection weights. As ten networks completed each training condition, their connection weights were averaged to produce a summary of how a typical network solved that particular problem. This summary is presented in Table 4-4.

Table 4-4

Operant perceptron connection weights for each cue type

		Problem and type			
		Four		Eight	
Unit Type		Corner	Wall	Corner	Wall
	Bias	-0.41	-0.40	-1.14	-1.14
Angle	90	-0.41*	0.00	0.00*	-1.14
	180	0.00	-0.40*	-1.14	0.00*
Color	Red	-0.66	-2.30	-0.57	-0.28
	Green	3.19*	-0.67	3.30*	-0.29
	Yellow	-2.29	-0.65	-2.18	-0.28
	Blue	-0.66	3.21*	-0.55	-0.29
	Orange	0.00	0.00	-0.29	-0.56
	Black	0.00	0.00	-0.28	-0.56
	Brown	0.00	0.00	-0.29	-2.18
	Purple	0.00	0.00	-0.28	3.30*
	White	0.00	0.00	0.00	0.00
Configuration	3 / 6	-1.31	0.00	0.00	0.00
	6 / 3	0.91*	0.00	0.00	0.00
	2 / 1	0.00	-1.32	-0.57	-1.12
	2 / 4	0.00	0.91*	-0.57	1.12*
	4 / 2	0.00	0.00	1.12*	-0.57
	1 / 2	0.00	0.00	-1.12	-0.57

**Corresponds to reinforced location*

An examination of this table reveals that, within the 4-location task, the bias and reinforced angle units assume negative values, while non-reinforced angle units assume a value of

0. This informs us that, before considering wall-length configuration or feature information, the network initially tends to turn off (output activity, and thus probability of investigating a location, approaching 0) at any given location. In the 8-location task, however, this is slightly different: while the bias remains negative, the reinforced angle assumes a 0 weight, while the non-reinforced angles assume a strong negative weight. Despite this difference, this pattern of weights, in absence of other cues, produces identical behaviour to the 4-location network.

It is only after the network considers other cues that it begins to overcome this negative association and develops a moderate probability of investigating a given location. Within the 4-location task, the wall-length configuration corresponding to the reinforced location assumes a positive value with magnitude slightly larger than the magnitude of the bias and the angle at that location. A similar result occurs in the 8-location task, where the correct wall-length configuration and the bias effectively cancel out and the correct angle has a weight of 0. In both of these cases, the net input is close to 0; the output unit's logistic function translates this into a 0.5 probability of acting, given those cues. In other words, for both the 4-location and 8-location task, if the networks encounter a location with the correct geometry but lacking any feature, they are as likely as not to choose to investigate that location. The overall choice frequency behaviour this produces will vary depending on the number of locations (see Tables 4-2 and 4-3), however, the underlying mechanism is identical. It is interesting to note that, ignoring features, the 'correct' wall-length configuration is reinforced on 50% of its presentations (the reinforced location and its rotational equivalent), while the 'incorrect' configurations present in any condition are reinforced 0% of the time, and the perceptrons' responses converge to match these probabilities. The operant perceptron has already been established to match probabilities in classical choice-behaviour tasks (Dawson et al., 2009); for it to exhibit this behaviour in a reorientation context reinforces Miller and Shettleworth's (2007) conceptualization of reorientation as an operant task.

The feature cue connection weights tell an unsurprising story in both the 4-location and 8-location task. The feature that was reinforced during training assumes a very strong positive weight, while the feature rotationally-opposite the reinforced location (i.e. the other location with identical geometric cues) assumes an equally-strong negative weight. The positive magnitude of the weight given to the correct feature far exceeds the negative value of the bias plus any incorrect cue – that is, the network has a high probability of acting when presented with the correct cue, even if both angle and wall-length configuration cues are incorrect.. Meanwhile, all other features assume a moderate negative weight. In the context of the geometric cues discussed above, this informs us that the network is inherently hesitant of investigating any location, but that the presence of a correct feature is sufficient to overcome this hesitancy.

Discussion

The operant perceptron's behaviour on these simulations allows us to generate novel empirical predictions. To begin, the network uses the same encoding for all conditions (4 locations or 8 locations, and wall reinforcement or corner reinforcement), and was able to converge in all of these conditions without difficulty with the same amount of training. Therefore, the operant perceptron predicts that similar mechanisms are at work regardless of the global shape of the arena, and that changing the number of locations of interest will have a negligible effect on the difficulty of the task. These predictions are broadly compatible with previous empirical work on multiple-location reorientation (Newcombe et al., 2010) but are incompatible with theories that include an implicit representation of the global environment (Miller & Shettleworth, 2007, 2008).

Furthermore, the operant perceptron does not predict any real difference between tasks where the locations of interest fall within corners, and tasks where such locations do not fall on corners. In both cases, networks were able to learn the task, encoding sufficient geometric cues to reorient and producing comparable behaviour when presented with transformed arenas. This behaviour persisted even if the cue types were completely novel, suggesting some degree of generalization – although the network predicts that the mechanism behind this generalization is inhibitory.

We can elaborate on this inhibitory mechanism by examining the connection weights in Table 4-4. Specifically, the networks learned that particular wall-length configurations signaled that a location was *not* reinforced, and learned that a particular color's rotational opposite was a reliable indicator of no reinforcement. When the networks were presented with the transformed arenas, they did not respond to the novel geometry at all – they had not learned that such configurations signaled no reinforcement. Instead, the network responds at chance values to each location, except for the two locations containing the ‘correct’ feature and its rotational opposite. Rather than developing an explanation of what the agent may be searching for in these cases, a study of connection weights informs us that we should instead be focusing on what the agent is *avoiding*. This tendency to emphasize excitation at the expense of inhibition when explaining learning is a tacit assumption present in many different theories of learning (Rescorla, 1967); the operant perceptron model reinforces this point, and reminds us of the need to check such assumptions.

Experiments

The operant perceptron has generated some interesting predictions on novel permutations of the reorientation task. Specifically, the operant perceptron makes two broad claims: first, that there is no appreciable difference in reorientation behaviour among groups trained with locations in corners or along walls, and second, that there is no appreciable change in difficulty when the number of salient locations changes. Do these predictions hold under laboratory conditions with live agents? To test these predictions, we conducted a series of basic reorientation experiments using human subjects.

Our experiments are organized into three studies. Study 1 involves two groups of subjects trained on a 4-location reorientation task; one group is trained on corner locations, and another trained on wall locations. The locations in these tasks correspond to Figures 4-1A and 4-1B.
Study 2 is analogous to Study 1, except that the training arenas have 8 locations of interest, as in Figure 4-1A.

Immediately after completing Study 1 or Study 2, each participant also completed the task described in the other study – Study 1 participants completed the 4-location task and then immediately progressed through the 8-location task exactly as described in Study 2, and vice versa. This allows a direct comparison of the difficulty of reorientation in arenas with 4 locations and 8 locations. Furthermore, this manipulation can test for order effects: did subjects learn either task faster, and did the first task facilitate learning the second? These comparisons are the focus of Study 3.

Study 1: Four Locations

Subjects

Participants were 36 University of Alberta undergraduates (30 female), who received course credit for participation. Recruitment criteria required participants to have normal color vision.

Apparatus

The environment was constructed using the fAARS-Lite platform (see Chapter 3 and Gutiérrez, 2012), which simulates first-person 3-D movement in a virtual world. This virtual world contained a number of rectangular arenas (17.2m by 8.6m), consisting of matte grey walls and floors with black, visually-obvious edges. The walls were high enough to extend beyond the subject's field of vision.

Subjects would arrive in an arena in its center, facing a random direction. Subjects could move their perspective through these arenas with the arrow keys.

Stimuli

Attention was called to locations of interest through brightly-colored cylinders (1.5m radius, same height as surrounding walls) placed against the walls of the arena. These locations were placed in two possible configurations, with the locations of interest being set at corners or along walls. These configurations correspond to Figure 4-1A and Figure 4-1B. During training, these cylinders had one of four colored textures placed over a white background: Green checkerboard, red diagonal stripes (upper-left to lower-right), yellow diamonds, or blue horizontal bars, listed in the order they appeared looking clockwise from the center. Figure 4-2 presents examples of these stimuli.

During testing, the positioning of these locations shifted into one of three possible configurations: Affine, Generalizing, or Featureless. In the Affine condition, each location had been shifted one "slot" clockwise: in Figure 4-1A, the green location had been present at Location 1 during training, and would be at Location 2 in an affine-transformed arena. In the Generalizing condition, the targets were shifted to a novel geometry – that of the other group's training condition. That is, a subject trained in Figure 4-1A would experience Figure 4-1B as its Generalizing condition. Finally, the Featureless transform removes all distinguishing information from the cylinders: in place of their brightly-colored patterns, subjects simply saw blank, white pillars.

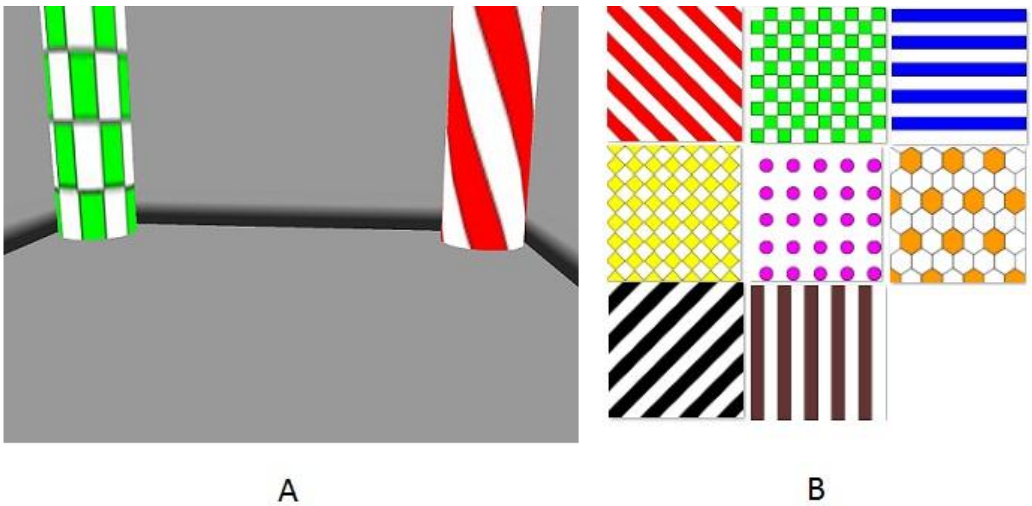

Figure 4-2: Example stimuli from the perspective of a subject in Study 1 (A), standing in the center of a rectangular arena with visually-salient locations of interest placed in the corners. These locations could have many possible color values (B), although some colors only appear in Study 2.

General Procedures

Participants were pseudo-randomly divided into four groups based on two possible categories (counterbalancing for gender). One of these divisions was based on target location: half the subjects would be trained with locations along walls, and half the subjects would be trained with locations at corners. Similarly, subjects were split into two reinforcement groups before the experiment began: those who would receive reinforcement at a location with a long wall on its left (group A) and those who would receive reinforcement at a location with a long wall on its right (group B).

Upon arrival, participants were instructed on how to move around in the virtual world, and were given time in a “welcome” room (A curved hallway with arrows pointing to a door at its end) to practice movement before the experiment began. Instructions were given to find a ‘correct location' inside each new room they saw; these instructions deliberately avoided the words "corner" and "wall". Participants were told that they made a choice by walking into a location, at which time they would see a display informing them whether their choice was correct or incorrect. Occasionally, they were told, the display would say "no feedback" regardless of the accuracy of their choice. To encourage a consistent strategy over time, subjects were told that they would be awarded points for correct choices (even if the display said "no feedback"), and that they should maximize their score.

Training

Training consisted of blocks of 10 presentations of the training arena. Each presentation allowed the subject any amount of time to move freely about the enclosure, but ended when they moved into one of the target locations, receiving feedback as described above before appearing in the center of the training enclosure again. Training continued until the participant made 8 "correct" choices in a single block, at which point they progressed into a non-reinforced training phase. During this phase, subjects had a 50% chance of seeing a "no feedback" message after making a choice. After making 8 correct choices during this phase, testing began.

Testing

Participants received five test trials in each of the three transformation conditions (Affine, Generalized, and Featureless). The test trials were presented in random order. Choices always received a "no feedback" response during this phase.

Following testing, subjects completed a post-testing retention test – 10 no-feedback trials in their original training enclosure. Subjects must have made 7 correct choices during this test to be included in analysis. Following this test, subjects were *not* immediately debriefed; instead, they proceeded to Study 3.

Results and Discussion

Five subjects failed the post-testing retention test, resulting in 31 subjects (26 female) included in analysis. Early analysis indicated that the data do not conform to normal distributions. Therefore, subjects' choice data were analyzed using randomization tests, employing bootstrapping methods to obtain confidence intervals on Cohen's d' measure of effect size (Edgington, 1995; Efron & Tibshirani, 1994). These methods are employed throughout the remainder of the discussion.

This analysis showed no difference in number of training blocks between the genders (males, M=2.0, SD=0.71; females, M=1.88, SD=1.34; d'=0.1, 95% CI [-0.56,1.32], p=0.86), nor between different reinforcement groups (group A, M=1.6, SD=0.63; group B, M=2.2, SD=1.6; d'=-0.48, 95% CI [-1.03,0.24], p=0.19). Similarly, subjects did not exhibit any significant differences in the amount of blocks required to learn the task if they were reinforced at walls or at corners (walls, M=2.25, SD=1.61; corners, M=1.53, SD=0.52; d'=0.60, 95% CI [0.02, 1.11], p=0.08).

Within the Affine test arena, there were no significant difference between groups in terms of the amount of choices made that were geometrically consistent (wall, M=0.22, SD=0.29; corner, M=0.09, SD=0.14; d'=0.58, 95% CI [-0.09,1.15], p=0.11) or featurally consistent (wall, M=0.74, SD=0.33; corner, M=0.89, SD=0.18; d'=-0.58, 95% CI [-1.18,0.03], p=0.11) with training. There were also no significant differences between groups in terms of featurally-consistent choices within the Generalized arena (wall, M=0.91, SD=0.16; corner, M=0.87, SD=0.25; d'=0.17, 95% CI [-0.53, 0.76], p=0.63). Finally, when subjects were tested in Featureless arenas, both groups made similar amounts of geometrically-consistent choices (wall, M=0.58, SD=0.35; corner, M=0.57, SD=0.34; d'=0.03, 95% CI [-0.59, 0.77], -p=0.96).

Since neither corner-reinforced nor wall-reinforced subjects showed any differences in choice behaviour, they were pooled together to test if their geometric or feature choices were significantly different from chance. Within the Affine arena, subjects' choices followed the feature at a rate significantly higher than chance (95% CI [0.73, 0.90], p<0.05), and made choices consistent with training geometry significantly less often than chance (95% CI [0.08,0.24], p<0.05). Feature-consistent choices were made in the Generalized arenas more often than chance (95% CI [0.82,0.95], p<0.05). Within the Featureless arena, however, subjects' choices did not significantly differ from chance (95% CI [0.46, 0.68], p>0.05).

Taken as a whole, these results lead us to two conclusions. First, ‘corner’ locations do not appear to be special in a reorientation context: subjects trained to visit wall locations produced behaviour statistically indistinguishable from the more classic corner-visiting group. Second, that behaviour suggests that subjects rely on features far more than geometry in this particular reorientation paradigm, to the point where they almost fail to encode geometry altogether. This result is consistent with other reorientation research which has shown that, in some conditions, feature cues can overshadow geometric cues (Bodily, Eastman, & Sturz, 2011; i.e. Horne & Pearce, 2009; Pearce et al., 2004), but such results are not at all universal.

Study 2: Eight Locations

Subjects

Participants were 35 University of Alberta undergraduates (22 female), who received course credit for participation. As in Study 1, participants were required to have normal color vision.

Apparatus

The apparatus was identical to that employed in Study 1, except in regards to stimuli.

Stimuli

The stimuli had the same general nature as in Study 1; only the possible color patterns were different. Four of the patterns were the same stimuli present in Study 1. The remaining four were brown vertical stripes, purple spots, black diagonal stripes (upper-right to lower-left) and orange hexagons. These stimuli were placed in the configuration depicted in Figure 4-1C.

As in Study 1, there were three transformed arenas as well. Transformation consisted of shuffling which color was present at a particular “slot”. An Affine transformation was created by shifting the colors two “slots” clockwise relative to training, which placed targets at novel wall configurations, but with the same angle (90° or 180°) relative to training. A separate, “Partial” transformation was created by shifting the colors one “slot” clockwise; in this condition, both the wall configuration and angle were different compared to training. Also as in Study 1, a featureless condition was included where all locations had identical pure-white colors in place of their original patterns.

General Procedures

The procedures followed were identical in all ways to Study 1, including the post-testing retention test and progression to Study 3 upon completion.

Results and Discussion

Five subjects failed the post-testing retention test, resulting in 30 subjects (19 female) included in analysis. No evidence of a gender effect (males, *M*=2.0, *SD*=0.45; females, *M*=1.95, *SD*=0.78; *d'*=0.08, 95% CI [-0.59, 0.94], p=0.85) nor an effect of reinforcement grouping (group A, *M*=2.1, *SD*=0.83; group B, *M*=1.8, *SD*=0.42; *d'*=0.51, 95% CI [-0.27, 1.21], *p*=0.28) was found among the number of blocks these subjects required to complete training. Similarly, no significant difference was found in the number of blocks needed for subjects to learn their reinforcement was at a corner or a wall (walls, *M*=2.06, *SD*=0.44; corners, *M*=1.85, *SD*=0.86; *d'*=0.30, 95% CI [-0.33,1.28], *p*=0.45).

In terms of choice consistency with particular cues, subjects exhibited a similar pattern to those in Study 1 in the Affine and Featureless conditions. Within the Affine arena, subjects in both the wall-target and corner-target conditions made similar proportions of choices consistent with geometry (wall, *M*=0.26, *SD*=0.43; corner, *M*=0.33, *SD*=0.41; *d'*=-0.17, 95% CI [-0.89,0.45], *p*=0.60) and consistent with features (wall, *M*=0.74, *SD*=0.43; corner, *M*=0.61, *SD*=0.45; *d'*=0.31, 95% CI [-0.36,1.19], *p*=0.376). Similarly, the Featureless arena saw subjects produce similar proportions of choices consistent with training geometry (wall, *M*=0.49, *SD*=0.38; corner, *M*=0.54, *SD*=0.45; *d'*=-0.12, 95% CI [-0.88, 0.55], *p*=0.71). The Partial condition, where both wall configuration and angle cues varied from training, also produced similar proportions of choices consistent with training geometry (wall, *M*=0.30, *SD*=0.42; corner, *M*=0.31, *SD*=0.40; *d'*=-0.03, 95% CI [-0.68,0.67], *p*=0.90) or with features (wall, *M*=0.7, *SD*=0.42; corner, *M*=0.65, *SD*=0.42; *d'*=0.12, 95% CI [-0.56,0.87], *p*=0.725).

In every case, subjects from both groups produced indistinguishable results, and were therefore pooled to test if their choices varied from chance. Within the Affine arena, subjects made choices consistent with the feature significantly more often than chance (95% CI [0.54,0.81], $p<0.05$), but did not make choices consistent with the original wall-length configurations more often than chance (95% CI [0.16,0.43], $p>0.05$). Within the Partial arena, subjects also followed the original wall-length configuration at chance rates (95% CI [0.17,0.45], $p>0.05$), and followed feature cues significantly more often than chance (95% CI [0.54,0.80], $p<0.05$). Finally, unlike in Study 1, subjects responded to the original wall-length configurations in the Featureless arenas significantly more often than chance (95% CI [0.39, 0.65], $p<0.05$).

In general, Study 2 supports Study 1's findings that walls are not significantly different from corners in terms of reorientation, even when the number of locations is increased. Features remain the best predictor of subject behaviour, and clearly dominate such behaviour when they are presented in conflict with geometry, regardless of whether that geometry is completely inconsistent with training (the Affine condition) or partially inconsistent with training (the Partial condition). However, unlike Study 1, subjects' behaviour in the Featureless arena clearly indicates that geometry was encoded during training. This is consistent with other literature which finds that humans are capable of completing reorientation tasks with more than four locations of interest (Newcombe et al., 2010).

This geometric overshadowing is the only noteworthy difference between Study 1's 4-location task and Study 2's 8-location task. While this type of overshadowing has been observed during reorientation, in some circumstances (Bodily et al., 2011; Pearce et al., 2004), this is usually not the case (i.e. Wall et al., 2004). It is quite possible that another version of the 4-location reorientation task in which overshadowing does not occur may find no difference between 4-location and 8-location reorientation, should it be repurposed for the latter. Testing this claim would require another experiment, beyond the scope of the current manuscript.

Study 3: Direct Comparison

Subjects

Subjects were participants from both Study 1 and Study 2, as described above. This included subjects who failed their original study's post-testing retention test.

General Procedures

After completing the post-testing retention test for their original study, subjects found themselves placed back in the "welcome" chamber, and were instructed that this marked the halfway point of the experiment. The protocol followed from here is identical to that described in Study 1, except with the stimuli set from whichever study the subject did not already see. Following completion of this second experiment, subjects were debriefed.

Results and Discussion

For discussions of the 'first task', subjects who passed their post-testing retention tests as described in Study 1 and Study 2 were included in analysis. For the 'second task', a separate post-testing retention test was performed (independent of the subject's test result in their original study), and subjects who performed poorer than 70% on this trial were excluded from analysis. Of the Study 1 participants, two failed this test, leaving 34 (28 female), while three of the Study 2 participants failed, leaving 34 (20 female) for analysis.

A large, and significant, order effect was discovered (initial study, M=1.94, SD=0.98; second study, M=1.36, SD=0.60; d'=0.7, 95% CI [0.41, 1.05], p=0.001), regardless of which task was completed first. For the first task, there was no significant difference in training blocks between the 4-location task and the 8-location task (4-location, M=1.90, SD=1.25; 8-location, M=1.97, SD=0.67; d'=-0.06, 95% CI [-0.77, 0.37], p=0.80). There was also no evidence of any significant differences in time taken to learn the second task (4-location, M=1.28, SD=0.46; 8-location, M=1.44, SD=0.70; d'=-0.27, 95% CI [-0.68, 0.18], p=0.27).

Two general conclusions can be drawn from these results. First, we do not find any evidence that increasing the number of locations makes a task harder to learn, both for naïve subjects and for subjects who have been trained in a different task. This conclusion is consistent with the operant perceptron's predictions, but inconsistent with models that adjust for the total number of locations of interest (such as Miller & Shettleworth, 2007, 2008). Second, this result combined with the order effect allows us to conclude that subjects learned the second task faster than the first task, regardless of the number of locations present in either task. This suggests that something facilitated the second task. Since this is a difference in the mean number of training blocks, as opposed to a difference in actual time taken to complete those training blocks, it cannot be an increase in familiarity with the virtual world.

One possibility is that, like the perceptron model discussed earlier, human subjects may be using a similar system to learn both tasks – a system in which the total number of locations is irrelevant. The perceptron accomplishes this by using an encoding that relies purely on local cues, where each location is considered independently of other locations during training. Alternative models based on matching current visual stimuli to previously-learned visual stimuli (i.e. Cheung, Stürzl, Zeil, & Cheng, 2008) have a similar property.

Model Evaluation

With both simulation results and experimental evidence in analogous tasks at hand, we return now to the operant perceptron, and evaluate where the patterns of behaviour agree and where they disagree. The operant perceptron made two broad classes of prediction, both of which withstood experimental scrutiny. That is, within both networks and humans, subjects learned 4-location reorientation and 8-location reorientation with equivalent amounts of training (Study 3), and within each task, no appreciable difference was found between wall locations and corner locations in terms of reorientation behaviour. In short, some interesting and possibly counterintuitive predictions that arose from the computer simulation were supported by the experimental studies that used human subjects.

Importantly, the computer simulations can also be used to derive additional predictions that can be compared to human performance. One example prediction is the pattern of response frequencies reported in Tables 4-2 and 4-3. These tables also report the response frequencies for human subjects from Study 1 and Study 2. Are the network's response frequencies appreciably different from the humans', or are they a plausible model?

To evaluate this, for each experimental condition, a 95% confidence interval was bootstrapped onto the mean of the human subjects' responses to each type of cue (Efron & Tibshirani, 1994). These confidence intervals were used above to compare the response rates to chance; here, they are used to compare the response rate to the mean network response frequency to the same cue type. As no difference was found in human subjects between wall groups and corner groups, the networks' corresponding locations response frequencies were averaged as well. These confidence intervals and their comparisons are given in Table 4-5.

Table 4-5

Comparing human and network response frequencies to specific cue types

		Response Rates			
		Geometric		Feature	
Task	Transform	Human (95% CI)	Network	Human (95% CI)	Network
Four	Affine	[0.08,0.24]	0.49	[0.73, 0.90]	0.50
	Generalized	N/A	N/A	[0.82,0.95]	0.57
	Featureless	[0.46, 0.68]	0.83	N/A	N/A
Eight	Affine	[0.16, 0.43]	0.44	[0.54, 0.81]	0.45
	Partial	[0.17, 0.45]	0.51	[0.54, 0.80]	0.36
	Featureless	[0.39, 0.65]	0.71	N/A	N/A

In all cases with 4 locations, the network model consistently predicted too many choices consistent with wall-length geometry, and too few choices consistent with features. This is not surprising, as humans consistently chose geometric-consistent locations in this task at a rate indistinguishable from chance, indicating that they did not encode geometric cues, while a review of the connection weights in Table 4-4 indicates that the networks did encode such cues.

In the 8-location task, the networks performed much closer to human behaviour in general, although again the networks tend to respond more frequently to geometry and less frequently to features than do humans. These discrepancies in response frequencies suggest a need for exploring alternative design decisions in the perceptron. In particular, one open question concerns how changing the encoding of cue patterns might affect perceptron responses, as well as relationships between the networks and the humans that are based upon response frequency measures.

General Discussion

The current manuscript explored reorientation in novel variations on the standard reorientation task, informed by a simple artificial neural network. Network simulations and experimental data allowed us to examine how behavior changes – or rather, does not change – when the locations of interest are placed at locations other than the corners of a quadrilateral arena. Furthermore, we examined this in the context of changing the number of salient targets from 4 to 8, and found that the difficulty in learning the task does not actually increase, and furthermore that skills learned in one task generalize to the other in both directions. These results were consistent within both simulation and experiment, and suggest that the behaviour governing reorientation involves processing the cues available at locations taken in isolation, regardless of the global structure of the environment. In other words, learning reorientation does not require comparing the current location to any other possible location, as is the case in Miller and Shettleworth's (2007, 2008) model.

These results – both simulated and experimental – suggest an interesting refinement to the hypothesis that angles are processed in a manner similar to features (Sturz, Forloines, & Bodily, 2012). Under this refinement, the location of interest serves as a visually-salient "focal point" – a reference from which wall-length and angle cues are determined. In a typical reorientation task, the corners of an arena create that focal point, but our results suggest that other visually-salient goals (here, pillars, but also possibly bowls of food, boxes with toys, and so forth) can create the same effect, even if they are not placed at corners. The angle the walls form at the location of interest therefore becomes a feature of that location, in much the same manner as traditional feature cues, such as color. Additionally, theories of viewpoint-matching (Cheung et al., 2008) propose that reorientation is largely a matter of learning broad visual stimuli when reinforced, then seeking to minimize the difference between one's current visual input and this learned image. This theory could also be capable of reorienting in arenas without corner-based locations. Interestingly, both this theory and the operant perceptron model learning as error-correcting based on a pattern of subject behavior, although they encode the available stimuli in dramatically different manners. It would be interesting to see how their predictions on choice frequencies differ, if at all.

The choice to use configuration unit encoding (i.e. a different unit turns on for each possible wall-length configuration) choice was made for consistency with existing literature (Horne & Pearce, 2010; Miller & Shettleworth, 2007, 2008) and earlier chapters (Chapter 1). If the simulation diverges from live-agent data, then this theoretical choice may not be appropriate. Indeed, a divergence appear in two aspects: the simulation encoded geometry in the 4-location task while human subjects did not, and the networks produced slightly different choice frequencies (Table 4-5). However, in spite of this, the model still correctly predicted several interesting results, such as the lack of a difference between corner-trained and wall-trained subjects and the similar difficulty of the 4-location and 8-location task. This suggests that the model needs adjustment, but that adjustment need not be extreme. This adjustment could take the form of parametric adjustment (Miller, 2009) or a change to how arena walls are encoded (i.e. thermometer coding, Dawson, Kelly, et al., 2010); these changes have important theoretical implications, if any prove more fruitful than raw wall-length configuration.

These developments provide examples of experimental results informing future modeling decisions, which is common practice in cognitive science. However, the current chapter, in contrast, demonstrated that modeling can quite easily inform experiment as well. A new experimental result or theoretical construct can revise an existing model, which in turn can be used to generate empirical claims in novel environments quickly and cheaply. If any of those predictions are of interest, future experimentation can be used to test these new hypotheses.

This methodological style - creating simple and plausible models which *behave*, and then generating hypotheses and experiments based on this behaviour – is an example of the synthetic approach to cognitive science (Dawson, 2004; Dawson, Dupuis, & Wilson, 2010). The synthetic approach can prove fruitful in breaking future deadlock or opening up novel research paradigms. For example, to the best of our knowledge, the current manuscript details the first attempt at systematically varying on the nature and number of locations during reorientation. The decision to investigate this comparison was motivated entirely by the structure of the neural network model.

This neural network model is a ripe avenue for future research. The operant perceptron successfully handled reorientation with assorted numbers of target locations, positioned at arbitrary points along the edge of an arena, using the encoding described above. This encoding can easily be extended, including extra input units to represent angles other than 90° or 180°, or to represent other sets of wall-length configurations than those used here. Putting these properties together allows this architecture to handle any polygonal arena with edge-defined locations of interest, including kites (Dawson, Kelly, et al., 2010; Pearce et al., 2004) and octagons (Newcombe et al., 2010). While we could extend the operant perceptron to see if it fits the data from some of the novel tasks (i.e. regular ocagons, Newcombe et al., 2010) in a manner similar to a more standard perceptron (Dawson, Kelly, et al., 2010), the synthetic approach would be to generate totally new predictions inspired by our findings. In this case, we might try non-uniform octagons (an arena type not yet investigated), or we might note other successes of the operant perceptron altogether, such as superconditioning (Chapter 1) or probability matching (Dawson et al., 2009), and branch out beyond reorientation into completely new paradigms.

Lewandowsky (1993) observed that computer modeling had its benefits, if done with care. The current manuscript illustrates all of these core ideas. A desire to increase mathematical rigor in the Miller and Shettleworth (2007, 2008) model led to the development of new tools – both the operant perceptron (Chapter 1) used in simulation and the fAARS-Lite platform (Chapter 3) used in data collection. These tools facilitated finding, and testing, the tacit assumption in reorientation literature that corners have some inherently special property. Finally, the simulation results indicating that there should be no difference in effort needed to learn reorientation in arenas with more locations fit the description for serendipitous findings in novel environments, a point emphasized by the same result appearing among human subjects. It would appear that, even after twenty years, Lewandowsky’s observations and advice for cognitive modellers still remains effective.

References

Bodily, K. D., Eastman, C. K., & Sturz, B. R. (2011). Neither by global nor local cues alone: evidence for a unified orientation process. *Animal cognition*, *14*, 665–674. doi:10.1007/s10071-011-0401-x

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, *23*(2), 149–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3742991

Cheng, K. (2008). Whither geometry? Troubles of the geometric module. *Trends in cognitive sciences*, *12*(9), 355–61. doi:10.1016/j.tics.2008.06.004

Cheng, K., & Gallistel, C. R. (2005). Shape parameters explain data from spatial transformations: comment on Pearce et al. (2004) and Tommasi & Polli (2004). *Journal of experimental psychology. Animal behavior processes*, *31*(2), 254–9; discussion 260–1. doi:10.1037/0097-7403.31.2.254

Cheng, K., & Newcombe, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic bulletin & review*, *12*(1), 1–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15945200

Cheung, A., Stürzl, W., Zeil, J., & Cheng, K. (2008). The information content of panoramic images II: View-based navigation in nonrectangular experimental arenas. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(1). doi:10.1037/0097-7403.34.1.15

Dawson, M. R. W. (2004). *Minds and Machines: Connectionism and Psychological Modeling*. Oxford, UK: Blackwell.

Dawson, M. R. W. (2008). *Connectionism and Classical Conditioning*. *Comparative Cognition and Behaviour Reviews* (Vol. 3, p. 115). Comparative Cognition Society. doi:10.3819/ccbr.2008.30008

Dawson, M. R. W., Dupuis, B., Spetch, M. L., & Kelly, D. M. (2009). Simple artificial neural networks that match probability and exploit and explore when confronting a multiarmed bandit. *IEEE Transactions in Neural Networks*, *20*(8), 1368–1371. doi:10.1109/TNN.2009.2025588

Dawson, M. R. W., Dupuis, B., & Wilson, M. (2010). *From Bricks to Brains: The Embodied Cognitive Science of LEGO Robots* (p. 331). Edmonton, AB: Athabasca University Press.

Dawson, M. R. W., Kelly, D. M., Spetch, M. L., & Dupuis, B. (2008). Learning about environmental geometry: a flaw in Miller and Shettleworth's (2007) operant model. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(3), 415–8. doi:10.1037/0097-7403.34.3.415

Dawson, M. R. W., Kelly, D. M., Spetch, M. L., & Dupuis, B. (2010). Using perceptrons to explore the reorientation task. *Cognition*, *114*(2), 207–26. doi:10.1016/j.cognition.2009.09.006

Edgington, E. S. (1995). *Randomization tests* (3rd ed.). New York, NY: Marcel Dekker, Inc.

Efron, B., & Tibshirani, R. J. (1994). *Introduction to the bootstrap*. New York, NY: Chapman & Hall.

Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, MA: MIT Press.

Gutiérrez, L. A. (2012). *The fAARS Platform For Augmented Alternate Reality Services and Games*. University of Alberta. Retrieved from https://era.library.ualberta.ca/public/view/item/uuid:b8d36cef-585d-4010-8124-9a22c1384087

Horne, M. R., & Pearce, J. M. (2009). Between-cue associations influence searching for a hidden goal in an environment with a distinctive shape. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*(1), 99–107. doi:10.1037/0097-7403.35.1.99

Horne, M. R., & Pearce, J. M. (2010). Conditioned inhibition and superconditioning in an environment with a distinctive shape. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(3), 381–94. doi:10.1037/a0017837

Lewandowsky, S. (1993). The rewards and hazards of computer simulations. *Psychological science*, *4*(4), 236–243.

Lubyk, D. M., & Spetch, M. L. (2012). Finding the best angle: pigeons (Columba livia) weight angular information more heavily than relative wall length in an open-field geometry task. *Animal cognition*, *15*(3), 305–12. doi:10.1007/s10071-011-0454-x

Miller, N. Y. (2009). Modeling the effects of enclosure size on geometry learning. *Behavioural processes*, *80*(3), 306–13. doi:10.1016/j.beproc.2008.12.011

Miller, N. Y., & Shettleworth, S. J. (2007). Learning about environmental geometry: An associative model. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(3), 191–212. doi:10.1037/0097-7403.33.3.191

Miller, N. Y., & Shettleworth, S. J. (2008). An associative model of geometry learning: a modified choice rule. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(3), 419–22. doi:10.1037/0097-7403.34.3.419

Newcombe, N. S., Ratliff, K. R., Shallcross, W. L., & Twyman, A. D. (2010). Young children's use of features to reorient is more than just associative: further evidence against a modular view of spatial processing. *Developmental science*, *13*(1), 213–20. doi:10.1111/j.1467-7687.2009.00877.x

Pearce, J. M., Good, M. A., Jones, P. M., & McGregor, A. (2004). Transfer of spatial behavior between different environments: implications for theories of spatial learning and for the role of the hippocampus in spatial learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(2), 135–47. doi:10.1037/0097-7403.30.2.135

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological review*, *74*(1), 71–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5341445

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, NY: Appleton-Century-Crofts.

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

Sturz, B. R., Forloines, M. R., & Bodily, K. D. (2012). Enclosure size and the use of local and global geometric cues for reorientation. *Psychonomic bulletin & review*, *19*(2), 270–6. doi:10.3758/s13423-011-0195-5

Tommasi, L., & Polli, C. (2004). Representation of two geometric features of the environment in the domestic chick (Gallus gallus). *Animal cognition*, *7*(1), 53–9. doi:10.1007/s10071-003-0182-y

Twyman, A. D., & Newcombe, N. S. (2010). Five reasons to doubt the existence of a geometric module. *Cognitive science*, *34*(7), 1315–56. doi:10.1111/j.1551-6709.2009.01081.x

Wall, P. L., Botly, L. C. P., Black, C. K., & Shettleworth, S. J. (2004). The geometric module in the rat: independence of shape and feature learning in a food finding task. *Learning & behavior*, *32*(3), 289–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15672824

Final Chapter

The goals of this thesis were to illustrate the various methodological components of interesting research style within cognitive science. These components were, in brief, formal, theoretical, improvisational, and synthetic. Each chapter had elements of these components, emphasizing them to a greater or lesser degree, in an endeavor to illustrate that, as a whole, productive research can follow.

Chapter 1 was motived initially by identifying a formal weakness in an influential model of associative learning. We investigated the theoretical consequences of this weakness, and found the model to be fatally flawed. This chapter also introduced the operant perceptron, a neural network model grounded in strong, well-established mathematical formalisms. Constructing this model also brought several tacit assumptions in reorientation to light – and questioning these assumptions became the motivation behind the rest of the thesis (Chapter 4 in particular).

Chapter 2 illustrated the theoretical consequences of adopting this new formalism, strengthening the link between perceptron models and established contingency theory. This stage is essential when adapting the perceptron to operant tasks, as operant learning amounts to estimating reinforcement contingencies. Such a stage is necessary: it is not sufficient to simply claim that because the perceptron proceeds according to the behaviour of an operant creature, the network is learning to accomplish operant tasks. The theoretical component here reinforces the need for rigor when developing new models – equations that don’t reflect reality don’t provide a true model.

Chapter 3 was a brief diversion discussing models, focusing instead on the development and deployment of a new virtual environment, fAARS-Lite. This platform was designed to simulate a wide array of classic spatial learning tasks, without requiring a great deal of computer expertise to employ. While this has natural applications in comparative cognition, in the context of this thesis, it served a secondary purpose: fAARS-Lite is an environment highly conductive to improvisation. Rather than design a specific tool for the simulations at hand, we developed a set of extremely basic elements which can be chained together to replicate an array of interesting tasks, many of which are discussed in the chapter. This platform is also freely available[6], in an effort to encourage others – particularly in comparative cognition – to attempt to adopt this research method. With a tool like fAARS-Lite at hand, such researchers may be inspired to create new paradigms within the virtual world, and then bring them to life with their animal subjects – a reversal of the usual order of comparative cognition experiments.

This reversal, where one allows model building to inform experimental design, is the final, synthetic component of the research method in practice. This is exemplified in Chapter 4, in which the operant perceptron's formalism led to questioning the assumptions underpinning modern reorientation research. The model generated several interesting predictions, as well as suggested mechanisms behind these effects. fAARS-Lite was used to empirically verify these findings, and in many cases, they withstood experimental scrutiny.

The thesis project as a whole also suggests several possible avenues for future research, in light of the four methodological elements. *Formally*, we could expand upon Chapter 2, exploring the formal links between the associative learning of the operant perceptron and other bodies of knowledge, such as information theory and Bayesian learning. *Theoretically*, we can explore alternate input encodings for the operant perceptron; each new encoding brings new assumptions and claims about the world which may or may not produce interesting results themselves. Alternatively, we could follow our experimental results from Chapter 4, particularly exploring how these results link in with the different perspectives currently under debate in reorientation literature. Perhaps we can *improvise* a new hypothesis out of which components sustain further scrutiny, and incorporate this into our model, which can then be used to *synthesize* a new theory to guide future research.

[6] fAARS-Lite is currently being prepared for release as open-source software; if you are interested, contact me for a trial.

Appendix 1: Technical Appendix to Chapter 1

In Chapter 1, we described a number of problems in the Miller-Shettleworth (2007, 2008) associative models. We described these problems purely in terms of the behaviour of the models. However, the source for this erroneous behaviour is ultimately mathematical, and emerges primarily from Miller and Shettleworth's choice to scale the Rescorla-Wagner (1972) equation by Equation 1-3 or Equation 1-4. This appendix will elaborate upon this, demonstrating why such scaling is incorrect from the perspective of calculus, and why such scaling produces incorrect results.

The main characteristic of the Miller and Shettleworth (2007, 2008) model is that it multiplies the Rescorla-Wagner (1972) equation by a probability term. The intention of this multiplication is to make the Rescorla-Wagner model operant in nature. However, it is this multiplication that causes the problems that were identified earlier in this paper. This multiplication also causes the model to make unexpected (and, we believe, unintentional) claims about time.

In Rescorla and Wagner (1972), the parameter β is explicitly defined as a learning rate parameter - a rate reflecting how much learning takes place within a given amount of time. β is held constant when the model is employed, because to do otherwise would "beg justification" (p. 82). A consequence of holding learning rate constant at each iteration of the model is that the model implies a constant amount of time passes during each sweep. However, because the learning rate β is held constant, this "time step" is usually suppressed when writing the equations.

We now express the Rescorla-Wagner equation in terms consistent with Miller and Shettleworth's (2007) approach, making time explicit:

$$\boldsymbol{\Delta V_i(t) = \alpha\beta(\lambda - \Sigma V_i(t_i))} \tag{a1-1}$$

Here, the subscript i refers to iteration: the change in weights from the current iteration to the next depends on the sum of the weights at the current time. In this equation, the learning rate parameter β is proportional to the amount of time that passes per iteration, $(\Delta t)/(\Delta i)$. If the equation is consistently applied to every cue at every iteration, then Δi=1, and thus β is proportional to Δt. Since Miller and Shettleworth always apply their equation at every iteration, this assumption holds. Therefore, by subsuming the constant of proportionality into α and setting $\beta=\Delta t$, we can substitute into Equation a1 and rearrange to form a ratio, as follows:

$$\boldsymbol{\Delta V_i(t) = \alpha\big(\lambda - \Sigma V_i(t_i)\big)\Delta t \rightarrow \frac{\Delta V_i(t)}{\Delta t} = \alpha(\lambda = \Sigma V_i(t_i))} \tag{a1-2}$$

By the definition of a limit, as the time step Δt approaches zero, this ratio will approach the instantaneous time derivative of associative strength, and thus:

$$\boldsymbol{\lim_{\Delta t \rightarrow 0} \frac{\Delta V_i(t)}{\Delta t} = \frac{\partial V_i(t)}{\partial t}} \tag{a1-3}$$

Because this equivalence holds, the model's associative strength will change at the same rate on each iteration (Equation a1-2) as it does at each unit of time (Equation a1-3).

In defining their models, Miller and Shettleworth (2007, 2008) multiply the entire Rescorla-Wagner equation by some probability term P_L. Both the 2007 (Equation 1-3) and 2008 (Equation 1-4) P_L terms are functions of associative strength at a particular iteration V_i, which is itself a function of time, and therefore Equation 1-3 and 1-4 are both also functions of time. Both P_L equations can be expressed more generally as,

$$\boldsymbol{P_i\big(V_i(t)\big) = (P \circ V)(t_i)} \tag{a1-4}$$

Simply multiplying this term into the Rescorla-Wagner equation, as Miller and Shettleworth (2007, 2008) have done in Equation 1-2, introduces a second time dependency to the system:

$$\boldsymbol{\alpha(\lambda - \Sigma V_i(t_i) * (P \circ V)(t_i) = \frac{\Delta V_i(t)}{\Delta t} * (P \circ V)(t_i)} \tag{a1-5}$$

If β=1 (and, as a consequence, Δt=1), this equation reduces to the right side of Equation 1-2, which Miller and Shettleworth label "ΔV". However, this is *not* the equivalent of ΔV if Δt=1, rendering Equation a1-3 invalid:

$$\boldsymbol{\lim_{\Delta t \rightarrow 0} [\frac{\Delta V_i(t)}{\Delta t} * (P \circ V)(t_i)] \neq \frac{\partial V_i(t)}{\partial t}} \tag{a1-6}$$

Rather, the proper time derivative would handle this second time dependency through applying the chain rule to Equation 1-a4 and accounting for the resulting $\Delta V/\Delta t$ term (the exact form of which would depend on whether Equation 1-3 or Equation 1-4 was used for P_L). Miller and Shettleworth (2007, 2008) did not do this, but continued to treat the composite Equation 1-2 as if it were supplying a proper change over time. Instead, the additional uncontrolled time dependency causes time in the simulation to flow at different rates for each location, depending on the current associative strength of that location.

This gives rise to the scaling problems discussed above: the scaling introduced by different P_L values is effectively changing the learning rate at each location on a given iteration, which results in different amounts of time passing between iterations at each location (including exactly zero time between iterations for net-negative locations in the M-S 2008 model, as P_L=0 in those cases). However, the model determines this scale by referencing weights at a given iteration, instead of after a given amount of time has passed - and due to the effective learning-rate scaling introduced by P_L, these are no longer equivalent considerations (Equation a1-6). In effect, this scaling gives rise to a situation where the change in weight at one iteration may alter the next moment for one cue, but some moment in the past for a different cue, allowing the future to influence the past. This is, of course, impossible.

For illustration, we return to the Wall et al. (2004) reorientation task described earlier, using Miller and Shettleworth's (2007, 2008) original parameters. Figure a1-1 directly compares both the M-S 2008 and M-S 2007 models' behaviour reported in terms of iteration with their behaviour reported in terms of time. Normally for Rescorla-Wagner models, the amount of time that passes at each iteration is defined by β, and scaled by the dimensionless salience term α; time passing per iteration is therefore proportional to $\alpha\beta$ (a constant). With Miller and Shettleworth's Equation 1-2, these are further scaled by the dimensionless P_L term, which varies over time. Therefore, in Figure a1-1, the "time" axis reflects the cumulative value of $\alpha\beta P_L$ at each iteration.

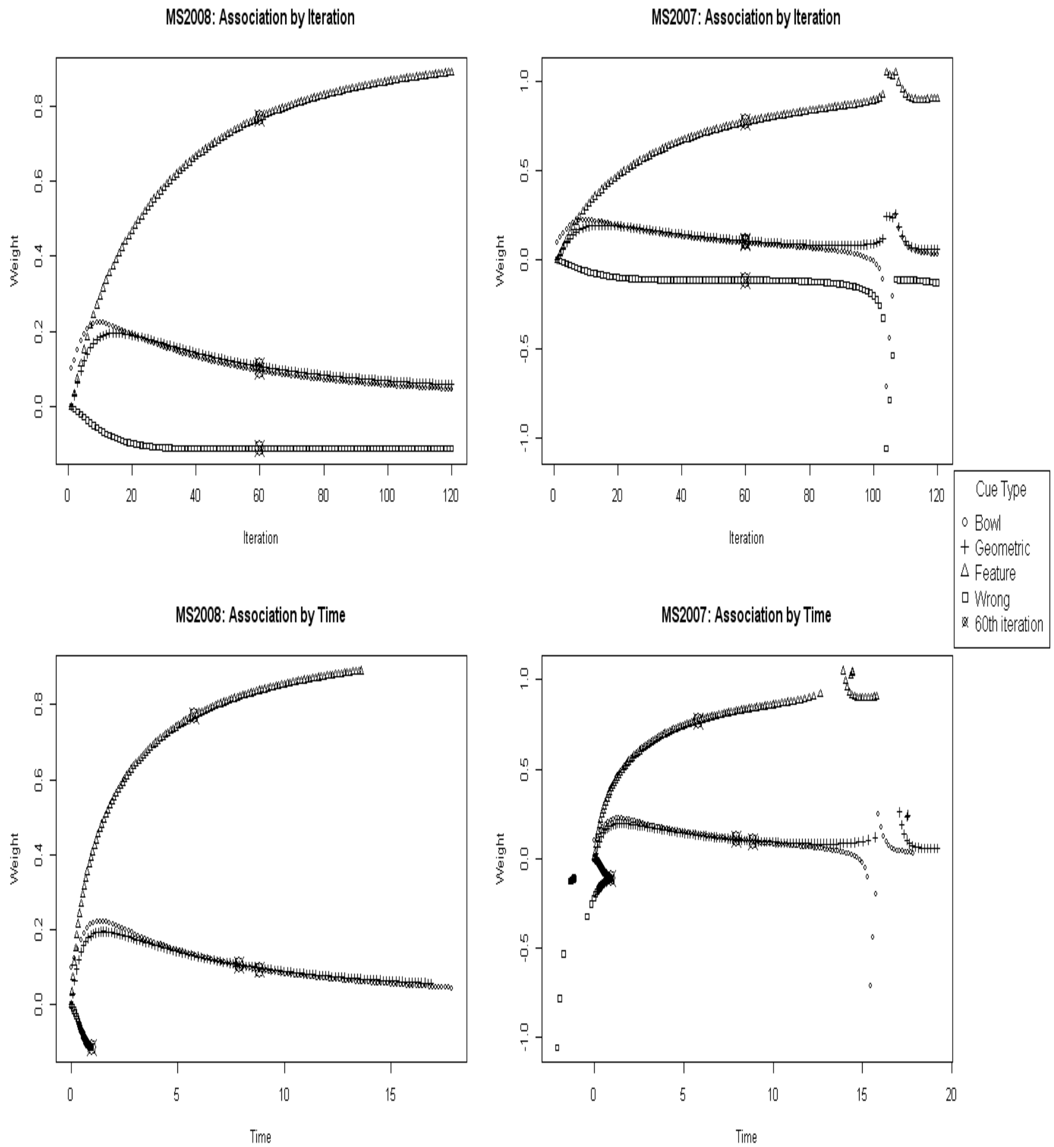

Figure a1-1. The Wall et al (2004) task, as interpreted by both the M-S 2008 (left) and M-S 2007 (right) models. The lower panels rescale the horizontal axis by $\alpha\beta P_L$, becoming proportional to time that has passed for each cue; the 60th iterations (considered to be simultaneous by the model) are highlighted.

Observe that that both models consider "simultaneous events" (such as the highlighted 60th iteration) which actually reflect different points in time. The nature of this asynchrony depends on the choice of equation for P_L. If Equation 1-3 is used, P_L is allowed to go negative for locations with sufficiently inhibitory net associative strength (Dawson, Kelly, Spetch, & Dupuis, 2008), therefore time begins to flow backwards. The singularities in the M-S 2007 model form immediately following when $t<0$ for the Wrong cue. (This corresponds to the exact point where ΣV_L changes signs in Equation 1-3.) The positiveness correction employed by the M-S 2008 model prevents time from flowing backwards, but does not prevent a given iteration from reflecting different points in time. In fact, when Equation 1-4 sets P_L to 0, Δt is also set to 0 for the corresponding cue (Wrong), such that every subsequent iteration reflects the same point in time.

In contrast, the operant perceptron model presented above does not suffer from this problem, as the underlying mathematics for updating its connection weights are formally equivalent to the Rescorla-Wagner equation (Dawson, 2008), and no scaling is applied. While the exact sequence of locations visited by the network may vary the amount of time the network spends at each location, this is functionally equivalent to adjusting the number of times each location is presented to the network - a course of action that does not introduce any uncontrolled time dependencies into the calculations for ΔV.

References

Dawson, M. R. W. (2008). *Connectionism and Classical Conditioning*. *Comparative Cognition and Behaviour Reviews* (Vol. 3, p. 115). Comparative Cognition Society. doi:10.3819/ccbr.2008.30008

Dawson, M. R. W., Kelly, D. M., Spetch, M. L., & Dupuis, B. (2008). Learning about environmental geometry: a flaw in Miller and Shettleworth's (2007) operant model. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(3), 415–8. doi:10.1037/0097-7403.34.3.415

Miller, N. Y., & Shettleworth, S. J. (2007). Learning about environmental geometry: An associative model. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(3), 191–212. doi:10.1037/0097-7403.33.3.191

Miller, N. Y., & Shettleworth, S. J. (2008). An associative model of geometry learning: a modified choice rule. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(3), 419–22. doi:10.1037/0097-7403.34.3.419

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, NY: Appleton-Century-Crofts.

Wall, P. L., Botly, L. C. P., Black, C. K., & Shettleworth, S. J. (2004). The geometric module in the rat: independence of shape and feature learning in a food finding task. *Learning & behavior*, *32*(3), 289–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15672824

Appendix 2: Operant and Standard Perceptron Equilibria

Chapter 2 discussed the established "standard" perceptron with logistic activation function (Dawson, 2004). Chapter 1 discussed a modified version of this perceptron, which employs an operant learning procedure. The purpose of this appendix is to demonstrate that these two perceptrons, as far as their equilibria (and thus their link to contingency theory) are concerned, are identical. This brief argument has three points: an existence proof, a behavioural justification, and an empirical demonstration.

First, observe that the operant perceptron's probability of acting in response to a particular pattern of stimuli is given by the logistic function of that pattern's net input. The logistic function (Equation 1-5) asymptotes at 0 or 1 for extremely positive or negative inputs – by definition, it is always nonzero and positive. Because of this, the perceptron will, over time, always have some chance of responding, even to consistently non-reinforced stimuli. Therefore, the perceptron convergence theorem (Rosenblatt, 1962) applies: for any problem that is linearly separable (see Dawson, 2008), the operant perceptron is guaranteed to find a solution in a finite number of sweeps. This is the same theorem that guarantees the convergence of the standard perceptron. The exact number of sweeps may vary between these models, but if a solution exists that one model can find, then the other will also find it.

Second, when one considers what the operant perceptron's choice behaviour entails, the fact that the two types of perceptron will find a solution to a linearly-separable problem is not surprising. The operant perceptron has some probability of randomly deciding to not investigate a particular pattern of cues: on these iterations, the model simply does not update the weights. This is akin to simply "skipping" a pattern during training. Perceptron training typically involves a random order of presentation for each pattern; if the frequency with which these patterns is allowed to vary, then we convert the perceptron into the operant perceptron. Critically, the mathematics behind the learning rules remains identical otherwise – the change between the two is analogous to a change in training set. We therefore expect an operant and a standard perceptron, trained on the same linearly-separable problem, to eventually reach the same equilibrium (but not necessarily in the same amount of time).

This claim has received experimental support. The operant perceptron was initially explored in the context of probability matching (Dawson, Dupuis, Spetch, & Kelly, 2009). In this study, the authors performed a direct comparison between the operant and the standard perceptrons over time, including their responses when reinforcement contingencies were changed mid-training. Figure a2-1 clearly shows that the two models followed slightly different paths, but nonetheless converged to the same equilibria.

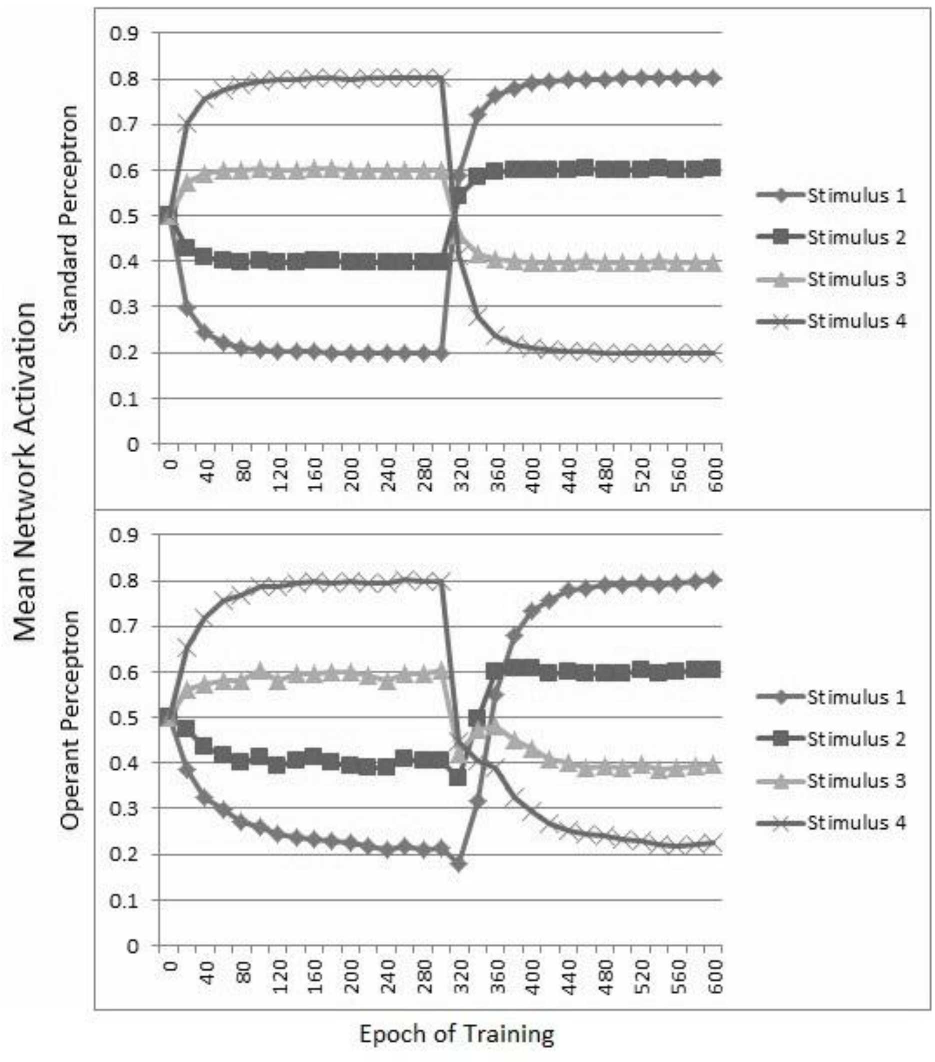

Figure a2-1: The equilibria for both the standard (top) and operant (bottom) perceptrons, allowed to train to convergence in a simple probability-matching task. Each stimulus is reinforced at a particular rate, and the perceptrons' equilibria matched that corresponding rate both before and after these rates were changed at epoch 300. The final equilibria were identical across both models. (Adapted from Dawson, Dupuis, Spetch, & Kelly, 2009; used with permission.)

In conclusion, Chapter 2's proofs of equivalence between some measures in contingency theory and the activity of a standard perceptron at equilibrium also apply to the operant perceptron. The exact path to equilibrium may vary, which has theoretical implications for certain psychological tasks, but here, only the equilibria themselves matter, and those do not vary.

References

Dawson, M. R. W. (2004). *Minds and Machines: Connectionism and Psychological Modeling*. Oxford, UK: Blackwell.

Dawson, M. R. W. (2008). *Connectionism and Classical Conditioning*. *Comparative Cognition and Behaviour Reviews* (Vol. 3, p. 115). Comparative Cognition Society. doi:10.3819/ccbr.2008.30008

Dawson, M. R. W., Dupuis, B., Spetch, M. L., & Kelly, D. M. (2009). Simple artificial neural networks that match probability and exploit and explore when confronting a multiarmed bandit. *IEEE Transactions in Neural Networks*, *20*(8), 1368–1371. doi:10.1109/TNN.2009.2025588

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