

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Spatial Factors in Social and Asocial Learning

Permalink

<https://escholarship.org/uc/item/0bf3r7bs>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 32(32)

ISSN

1069-7977

Authors

Metz, Alex
Shultz, Thomas

Publication Date

2010

Peer reviewed

Spatial Factors in Social and Asocial Learning

Alexander Metz (alexander.metz@mail.mcgill.ca)

Department of Psychology, McGill University, 1205 Avenue Docteur Penfield
Montreal, QC H3A 1B1 Canada

Thomas R. Shultz (thomas.shultz@mcgill.ca)

Department of Psychology and School of Computer Science, McGill University, 1205 Avenue Docteur Penfield
Montreal, QC H3A 1B1 Canada

Abstract

Asocial learning is a mechanism by which innovations develop, and social learning is a mechanism by which innovations spread. Penetration of an innovative behavior through a population is measured by the proportion of the population that possesses the innovation. Via agent-based modeling, we examine innovation diffusion with agents learning and interacting in space. Simulations show that innovation spread systematically deviates from differential equations of the proportion of the population that has the innovation. Mediation analysis confirms that boundary surface length of groups having the innovation accounts for these spatial effects. Proportion of asocial innovative learners increases surface length which, in turn, increases social imitative learning.

Keywords: Social learning; asocial learning; imitation; innovation; spatial simulation; surface length; mediation analysis; agent-based modeling.

Introduction

Since Darwin's theory of evolution, researchers have sought to understand how organisms adapt to their environment to maximize their reproductive potential. In addition to biological evolution, some lasting adaptations manifest themselves through animal phenotypes with no genetic changes. Innovative behaviors allow relatively quick adaptation to rapidly changing environments, and can spread and persevere (Laland, Boyd, & Richerson, 1996; Reader & Laland, 2003). For our purposes, *innovation* refers to an adaptive behavior pattern with relative novelty.

Innovative behaviors can be acquired through either asocial or social learning. In asocial learning, an innovation is acquired through individual experience. In social learning, an innovation is acquired from a conspecific demonstrator (Heyes, 1994). Because individual discoveries are costly, they may occur in only a few key individuals through asocial learning and then diffuse through a population via social learning (Laland, Boyd, & Richerson, 1996). The dynamics of innovation diffusion and imitation are an important aspect of collective social cognition and behavior, and thus relevant to integrative cognitive science (Goldstone & Gureckis, 2009).

However, patterns of sequential spread in a population resembling those predicted by some models of social learning can result from asocial processes alone. It is not always clear which type of learning drives innovative acquisitions, so it is thus important to disentangle social

from asocial learning. One useful method is diffusion curve analysis, or DCA (Reader, 2004; Franz & Nunn, 2009). Diffusion is the change in frequency of an innovative behavior in a population over time. In DCA, the shape of the curve is used to infer whether social or asocial learning is the mechanism of diffusion.

Our purpose is to simulate the spatial diffusion of innovation and compare the results to DCA predictions. Although there is a rich literature on learning in laboratory experiments, understanding of how social learning occurs in the wild is limited.

S-shaped logistic curves are predicted by DCA to characterize social learning. If the amount of social learning at a given time step is proportional to both the number of possible demonstrators and the number of possible learners, then it can be obtained as the product of the proportion of the population that knows the innovation and the proportion of the population that does not know the innovation (Laland, Boyd, & Richerson, 1996). This corresponds mathematically to the differential equation:

$$\Delta u = R_S u(1 - u) \quad (1)$$

where R_S is a constant rate of social learning, and u is the variable proportion of the population with the innovation.

In a population of only innovators, assuming no social learning, the following differential equation applies:

$$\Delta u = R_I (1 - u) \quad (2)$$

where R_I is the rate of innovation. As more innovators learn, the number of naïve innovators decreases in a decelerating curve (Franz & Nunn, 2009).

Social and asocial learning are not mutually exclusive. In an analysis of data from research by Hinde and Fisher (1949) on innovation spread in birds, Lefebvre (1995) concludes that milk-bottle-opening likely spread by some form of social learning from many unique points of origin. This is supported by evidence that some birds open bottles spontaneously without any prior experience with bottles or demonstrators (Sherry & Galef, 1984). Thus, asocial learning can occur alongside social learning, and Equations 1 and 2 can be summed to accommodate this:

$$\Delta u = R_S u(1 - u) + R_I (1 - u) \quad (3)$$

Equation 3, however, applies only to a population where every member is capable of being an innovator and a social learner. It may be more realistic to assume that only a certain proportion of the population is capable of either of

these things. No explanatory power is lost in making this assumption as these proportions can be set to 1, and the resulting model is only slightly less parsimonious. To accommodate this variation in ability, Equation 3 can be modified by multiplying the innovation and social learning parts of the equation by their corresponding proportions, I and S , respectively:

$$\Delta u = SR_s u(1-u) + IR_I(1-u) \quad (4)$$

We refer to Equation 4 as the DCA equation. Based on the proportion of social learning compared to asocial learning, this differential equation generates a curve with either a logistic shape (greater social learning) or a decelerating shape (greater asocial learning). The DCA equation has been applied in various experimental contexts, including the diffusion of innovations in humans from peer and media influences (Lekvall & Wahlbin, 1973) and bystander effects in the diffusion of foraging techniques in pigeons (Laland, Boyd, & Richerson, 1996).

The DCA equation relies on one key variable: the proportion of the population that knows the innovation. This proportion thus serves as both the dependent and independent variable in the differential equation. Here we test the results of spatial simulations against the predictions of the DCA equation. The diffusion of innovation is in part a spatial process, a fact captured by the simulations, but not by the DCA equation. We answer several questions. What are the essential differences between asocial and social learning and how can these two types of learning be identified in wild populations? Does the DCA equation account for all aspects of these issues, or are other approaches required? Are these features realistic, or are they artifacts of abstract simulations?

To explore the spatial diffusion of an innovation, our simulations create a two-dimensional space containing agents. Depending on their genotype, agents can be innovators and/or social learners. Parameters of the simulation include the proportions of innovators and social learners, just as the DCA equation uses these factors as variables. Comparing the rate of learning in the simulation to the rate of learning predicted by the DCA equation could provide insight into any potential spatial factors affecting innovation diffusion.

Methods

The simulation is set on a torus, a 25 by 25 lattice in which each edge touches the opposite edge. Each of the 625 tiles contains one agent with on/off genes for innovation and social learning.

Agents with an activated innovation allele can spontaneously discover the innovation at a fixed innovation rate of .025. Agents with an activated social-learning allele can copy the innovation from their neighbors: for every adjacent neighbor that knows the innovation, a social learner's chance of learning the innovation increases by .25. The ten-fold difference between the success of social and asocial learning is based on an assumption of differential learning costs: if asocial learning has a greater cost and

requires more resources than social learning, it should occur at a slower rate than social learning. The simulation experiment assigns genes to individual agents probabilistically depending on the proportion of social and asocial learners specified in simulation parameters. The simulation runs for 80 learning cycles, recording agent behavior, the times at which agents learn, and the neighbors from whom they learn if the learning is social.

The effect of number of innovators was investigated in simulations with the proportion of innovators ranging from .05 to 1.0, holding the proportion of social learners at 1.0. The effect of number of social learners was studied with simulations varying proportion of social learners ranging from 0 to 1.0, holding the proportion of innovators at 1.0.

Results

Figures 1-6 plot the change in the proportion of the population that knows the innovation over time, averaged across five runs. Figures 1 and 2 depict the results from varying the proportion of the population with the asocial learning allele when the whole population has the social learning allele. Figure 1 shows predictions of the DCA equation, and Figure 2 presents simulation results.

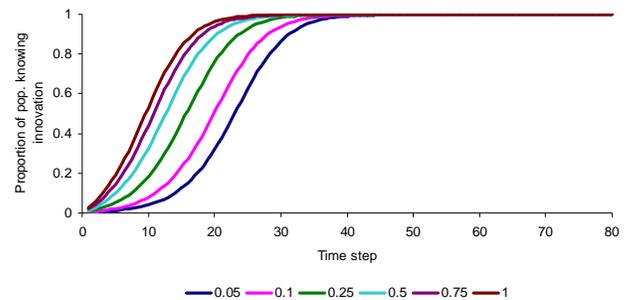


Figure 1: Diffusion curves predicted by DCA equation with asocial learning rate = .025, social learning rate = .25, proportion of social learners = 1, and the proportion of asocial learners varying from .05 to 1.

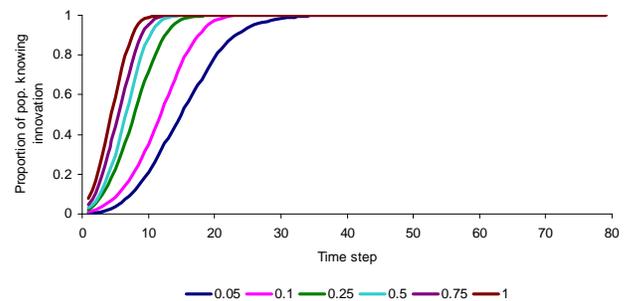


Figure 2: Simulations with asocial learning rate = .025, social learning rate = .25, proportion of social learners = 1, and proportion of asocial learners varying from .05 to 1.

These results reveal subtle but noticeable differences between the DCA equation and the simulations. For Figures 1 and 2, the whole population is capable of social learning;

what changes across curves is the proportion of the population capable of asocial learning. In Figure 1, the curves determined by the DCA equation appear more parallel than they do in the simulation results of Figure 2.

We can understand these differences by considering the DCA equation itself. This equation (4) has a social learning component (left half) and an asocial learning component (right half). Recall that the DCA equation's key variable is the proportion of the population that knows the innovation. At the beginning, the innovative behavior is introduced into the population by asocial learning, so the proportion of the population that can do asocial learning has a large effect as seen in Figure 1. Because this proportion of asocial learners is different in every curve, the curves differentiate quickly. However, as the proportion of the population that knows the innovation increases, the social learning component of the DCA equation has a greater effect. Because all of the curves in Figure 1 have the same social learning settings, with the proportion of social learners S set to 1 and the rate of social learning R_S set to .25, their learning rates are very similar after this original differentiation, causing the observed parallelism. Thus, the parallel nature of the equation-produced curves in Figure 1 is a direct consequence of using the proportion of the population that knows the innovation as the key independent variable.

The lack of parallelism in simulation curves can be quantified by examining the maximum learning slope for each curve, which represents the amount of learning when u , the proportion of the population that knows the innovation, equals .5. This is the point that maximizes the product $u(1-u)$ and thus also maximizes innovation spread according to the DCA equation. Figure 3 presents mean maximum slopes of diffusion curves as a function of the proportion of the population with the innovation allele.

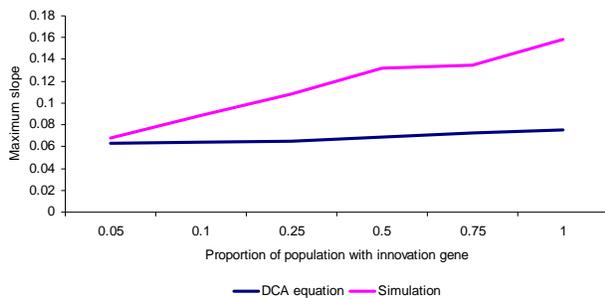


Figure 3: Maximum slope of curves (where $u = .5$) from the DCA equation and the simulations.

Figure 3 indicates that the maximum slope of each curve from the DCA equation is relatively stable across variation in number of innovators, consistent with a constant social learning component in the DCA equation. The corresponding simulations, however, do not follow this pattern; rather than being stable, the maximum slope increases with the proportion of innovators.

As Figure 6 indicates, there is no discrepancy between the asocial learning component of the DCA equation and

asocial learning in simulations. Thus we can infer that this increase in maximum slope across number of innovators is due to social learning. This implies that increasing the proportion of the population with the asocial learning allele speeds innovation spread in the simulation, which is exactly what we see in Figures 1 and 2.

Analogously, Figures 4 and 5 depict results from adjusting the proportion of the population with the social learning allele when the whole population has the asocial learning allele. Figure 4 shows predictions of the DCA equation while Figure 5 presents simulation results. Again, the curves produced from the simulations have a greater maximum learning slope than the curves predicted by the DCA equation, and these discrepancies increase with the proportion of the population that is capable of social learning.

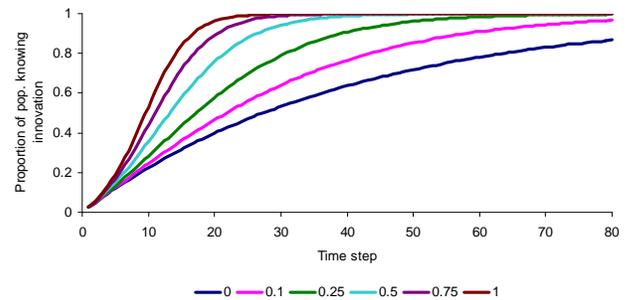


Figure 4: Diffusion curves predicted by DCA equation with asocial learning rate = .025, social learning rate = .25, and proportion of social learners varying from 0 to 1.

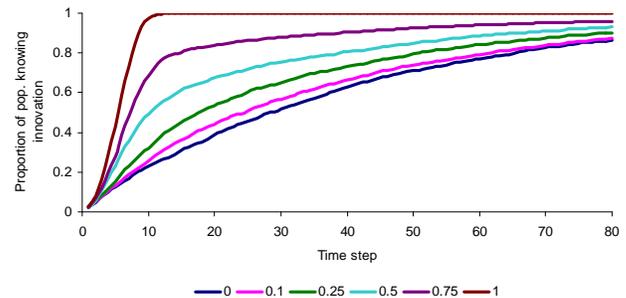


Figure 5: Simulations with asocial learning rate = .025, social learning rate = .25, proportion of social learners varying from 0 to 1, and proportion of asocial learners = 1.

With purely asocial learning ($S = 0$), the DCA equation closely tracks simulation results. The absolute differences between the equation and simulations averaged below .01 across all time steps. The lowest navy blue curves in Figures 4 and 5 are nearly identical. These two curves are re-plotted in Figure 6 to emphasize the overlap. This is the only simulation curve that the DCA equation successfully predicts. This predictive success makes sense because asocial learning in the simulation occurs as a random event based on a fixed probability, just as in the equation. Therefore, discrepancies between all other DCA and

simulation curves must result from social learning or possible interactions between social and asocial learning.

A possible cause of the increase in social learning as the proportion of innovators increases (Figure 3) is boundary surface length, the length of the perimeter surrounding groups of agents that know the innovation. These boundaries mark the area where naïve agents can learn the innovation. Thus, increasing this area should increase the speed of innovation spread.

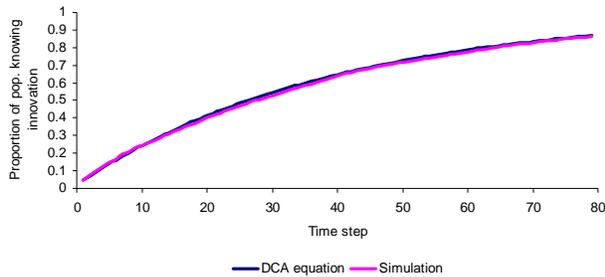


Figure 6: Diffusion curve predicted by the DCA equation compared to simulation results. Asocial learning rate = .025 and proportion of asocial learners = 1.

According to this analysis of the simulations, the spatial distribution of the agents that know the innovation affects social learning. Consider each innovator as a start point for an island of social learners. There will be more such islands when there are more initial innovators. More innovation islands generate more surface length and therefore more

social learning. This suggests an interaction effect with the proportion of innovators in the population: when there are multiple initial innovators, there is a higher likelihood that more social learning will occur as a result of greater surface length. When there are fewer initial innovators, less social learning will occur as a result of less surface length.

Figure 7 shows two plots from simulations exemplifying this argument. These two tori present simulation outputs, each depicting the point where one-half the population possesses the innovation. In 7A, where the proportion of innovators = .05, there are two islands, resulting from a few early innovators. In 7B, where the proportion of innovators = 1, there are upwards of nine islands due to more innovators. Although the proportion of the population possessing the innovation is the same in both worlds, surface length is much greater for the simulation that was initialized with a higher proportion of innovators.

Thus, an explanation for the discrepancies between the predictions of the DCA equation and the simulation results is that asocial learning increases the number of start points for social learning, and therefore the emerging amount of surface length. Because surface length determines the amount of social learning that can take place, social learning and innovation spread increase substantially as surface length increases. Thus, increasing asocial learning increases social learning in the simulation (Figure 2) but not in the DCA equation (Figure 1; see Figure 3 for direct comparison). This explanation can be further validated by a mediation analysis (MacKinnon et al, 2007).

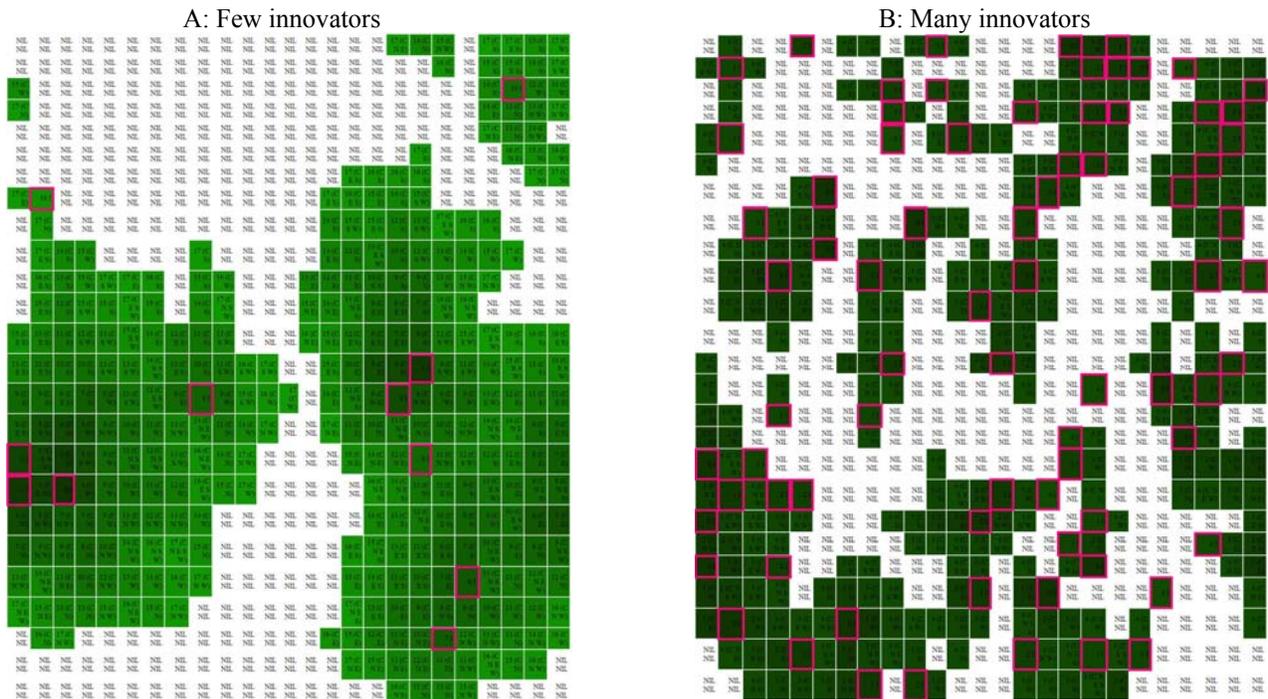


Figure 7: Two worlds with innovator proportions of .05 (A) and 1.0 (B). Time of acquisition is indicated by color saturation. Innovators are outlined in pink.

Mediation analysis is a type of linear regression that evaluates the relative effects of an independent variable (here, proportion of asocial learners) and a mediating variable (here, surface length) on a dependent variable (here, amount of social learning). The idea is that the independent variable affects the dependent variable, not only directly, but also indirectly via a mediating variable.

A mediation analysis of the simulation data across the six increasing proportions of asocial learners shows that 90.2 percent of the variance in the amount of social learning caused by variance in the proportion of asocial learners is mediated by surface length (total effect = 19.233 [$\beta = .945$], mediation effect = 17.351 [$\beta = .912 * .934 = .852$], $p < 0.0001$). As shown in Figure 8, the direct effect of the proportion of asocial learners on the amount of social learning becomes non-significant after controlling for the mediating variable of surface length, implying full mediation. This mediation analysis lends statistical support to the idea that surface length is the mechanism through which asocial learning causes social learning to speed up.

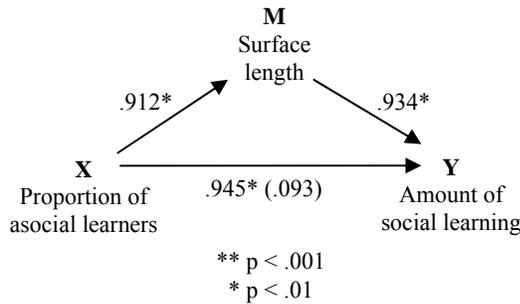


Figure 8: Standardized regression (beta) coefficients for mediational analysis. The path from X to Y falls to non-significance after controlling for the mediating variable of surface length (as indicated by the small coefficient in parentheses).

In summary, the simulations indicate that increasing social learning by adding more asocial learners increases surface length, and therefore increases the speed of social learning. This goes beyond the DCA equation which takes only the proportion of the population that knows the innovation as its independent variable. Also, speed of innovation spread is reduced as the number of agents with the social learning allele decreases.

Discussion

Our results show a difference between the DCA equation and the simulations, and this difference derives from the spatial factor of surface length. In the simulation, surface length is causally related to both social and asocial learning. Asocial learning increases surface length and surface length, in turn, increases social learning. The DCA equation, whose only independent variable is the proportion of the population that knows the innovation, does not capture this spatial factor. It is possible that the DCA equation could be

improved on by a more sophisticated mathematical model that incorporates surface length.

A fundamental question is whether or not these results apply to the real world. After all, the results are a consequence of the design of the simulations. There is a high viscosity in the design, meaning that agents can only learn from their directly adjacent neighbors. This characteristic is presumably the cause of the spatial effect. If an agent could learn from any other randomly-selected agent, then the spatial arrangement of agents would have no bearing on the results. Therefore, the results are only applicable to real-world scenarios where social learning depends highly on spatial proximity. With tools like the telephone and internet, which allow social learning to take place across oceans, these results may not apply to diffusion of innovation for many human populations. This is not to say that diffusion in humans is random, but rather that these present simulations may be too constrained to model it. However, the current results do seem applicable to populations where social learning is heavily dependent on proximity, which would include a lot of human learning based on face-to-face interactions.

This consideration points to a distinction between *geographic* and *social-network* analysis. The simulations we present here are examples of geographic analysis, with agents learning from their immediate neighbors. Social networks can transcend spatial proximity by using communication technologies to cover great distances. This difference is not just one of viscosity but also of structural complexity, because social networks are often more complicated than geographic relationships.

Franz and Nunn (2009) developed a method of social network analysis called network-based diffusion analysis, or NBDA. NBDA uses the social network of a population and the times at which they learn innovations to probabilistically determine whether the learning mechanism is social or asocial. Their method of social network analysis seems promising, although it requires the researcher to determine the social network of a population. Such specification may not be feasible in excessively large populations. There are also cases where a geographic analysis may be more appropriate because some environments are in fact viscous (e.g., Lefebvre, 1995).

Also, Franz and Nunn's main interest was in detecting social or asocial learning when one such learning method was exclusively present. In contrast, our research used various, systematic combinations of these two learning mechanisms. Model sensitivity to such combinations of social and asocial learning is more interesting and important than detection of pure cases. Studying such combinations is critical to discovering interactions between social and asocial learning, as highlighted in our results.

A lattice structure permitting interaction only between immediately adjacent neighbors is actually a special case of a network that provides only those links (or edges). Thus, a generalization of our results would entail testing whether an analog of surface length would facilitate information

diffusion in networks of various topologies. Such an analogy might be the number of directed links between agents who possess, and agents who lack, an innovation. If such links indicate direction of causal influence, then it would be important to count the links from knowledgeable to naive agents; if links indicate friendship choice, then count the links from naive to knowledgeable agents, because agents are likely to be influenced by those they consider to be friends.

The original aim of this project was to look for ways to disentangle social learning from asocial learning through a spatial analysis of the diffusion of innovative behavior. The results suggest that a greater proportion of asocial learners results in more innovation islands and greater surface length. Although it may be difficult to determine surface length in wild populations, counting islands in a topographic analysis of observations of innovative behavior would seem feasible.

The spatial effect of surface length provides a mechanism to disentangle social and asocial learning that is not available in diffusion curve analysis. This kind of spatial analysis could become another valuable tool to measure and understand the differences between social and asocial learning. One next step is to apply the ideas developed from this simulation to real biological data. In doing so, we may be able to contribute new understanding of how adaptive innovations spread and how they interact with evolution. Another planned thrust is to study how evolution selects the best proportions of social and asocial learning alleles under different environmental conditions (Laland et al., 1996; Shultz, Hartshorn, & Hammond, 2008; Shultz, Hartshorn, & Kaznatcheev, 2009). In such research, faster learning cycles can be nested within slower evolutionary cycles.

Acknowledgments

This research is supported by a fellowship from the McGill Faculty of Science to AM, and operating grants to TRS from the Natural Sciences and Engineering Research Council of Canada and a Dean's Excellence Award from the McGill Faculty of Science. We are grateful to Simon Reader, Louis Lefebvre, Artem Kaznatcheev, Stuart Wright, and Charles Garfinkle for ideas and insights on our project.

References

- Fisher, J., & Hinde, R. A. (1949). The opening of milk bottles by birds. *British Birds*, 42, 347-357.
- Franz, M., & Nunn, C. L. (2009). Network-based diffusion analysis: A new method for detecting social learning. *Proceedings of the Royal Society B*, 276 (1663), 1829-1836.
- Goldstone, R. L., & Gureckis, T. M. (2009). Collective behavior. *Topics in Cognitive Science*, 1 (3), 412-438.
- Heyes, C. M. (1994). Social learning in animals: Categories and mechanisms. *Biological Reviews*, 69, 207-231.
- Laland, K. N., Boyd, R., & Richerson, P. J. (1996). Developing a theory of animal social learning. In C. M.

- Heyes & B. G. Galef (Eds.), *Social learning in animals: The roots of culture*. San Diego: Academic Press.
- Lefebvre, L. (1995). The opening of milk bottles by birds: Evidence for accelerating learning rates, but against the wave-of-advance model of cultural transmission. *Behavioural Processes*, 34, 43-54.
- Lekvall, P., & Wahlbin, C. (1973). A study of some assumptions underlying innovation diffusion functions. *The Swedish Journal of Economics*, 75 (4), 362-377.
- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, 58, 593-614.
- Reader, S. M. (2004). Distinguishing social and asocial learning using diffusion dynamics. *Learning and Behavior*, 32 (1), 90-104.
- Reader, S. M., & Laland, K. N. (Eds.). (2003). *Animal innovation*. Oxford: Oxford University Press.
- Sherry, D. F., & Galef, B. G. Jr. (1984). Cultural transmission without imitation: milk bottle opening by birds. *Animal Behaviour*, 32, 937-938.
- Shultz, T. R., Hartshorn, M., & Hammond, R. A. (2008). Stages in the evolution of ethnocentrism. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1244-1249). Austin, TX: Cognitive Science Society.
- Shultz, T. R., Hartshorn, M., & Kaznatcheev, A. (2009). Why is ethnocentrism more common than humanitarianism? In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2100-2105). Austin, TX: Cognitive Science Society.