

Plasticity and language: an example of the Baldwin effect?

Kevin Zollman

Department of Philosophy
Baker Hall 135
Carnegie Mellon University
Pittsburgh, PA 15213-3890

Rory Smead*

Department of Logic and Philosophy
3151 Social Science Plaza A
University of California
Irvine, CA 92697

October 20, 2009

Abstract

In recent years, many scholars have suggested that the Baldwin effect may play an important role in the evolution of language. However, the Baldwin effect is a multifaceted and controversial process and the assessment of its connection with language is difficult without a formal model. This paper provides a first step in this direction. We examine a game-theoretic model of the interaction between plasticity (represented by Herrnstein reinforcement learning) and evolution in the context of a simple language game. Additionally, we describe three distinct aspects of the Baldwin effect: the Simpson-Baldwin effect, the Baldwin expediting effect and the Baldwin optimizing effect. We find that a simple model of the evolution of language lends theoretical plausibility to the existence of the Simpson-Baldwin and the Baldwin optimizing effects in this arena, but not the Baldwin expediting effect.

Humans and many other animals are extremely adaptive. Individuals are capable of changing their behavior in response to a wide variety of different

*The authors would like to thank Brian Skyrms, Bill Harms, Kyle Stanford, Brad Armendt, and Santiago Amaya for helpful suggestions.

environmental circumstances. This phenotypic plasticity has undoubtedly been shaped by natural selection, and its presence has potentially influenced the evolution of other less plastic traits. The relationship between natural selection, adaptation, and phenotypic plasticity have been the subject of significant discussion prompted by a multifaceted process proposed independently in the late 19th century by Lloyd Morgan, J. Mark Baldwin and Fairfield Osborne called the “Baldwin effect.” The Baldwin effect is an interaction between phenotypic plasticity and selection by which the presence of plasticity can alter the course of evolution and acquired characters can become genetically assimilated.

The discussion of the different aspects of the Baldwin effect has primarily focused on the theoretical plausibility of these components. Although far from decided, there has been significant negative criticism which regards the effect as unimportant, unlikely, or irrelevant. Some recent models have suggested that at least parts of the Baldwin effect may be more common than previously presumed (Ancel, 1999; Smead and Zollman, 2008).

Many studies of the Baldwin effect have focused on individual creatures adapting to a changing external environment. The changing environment creates a circumstance where there is some advantage to plasticity, since flexible individuals can respond to a new environment better than inflexible ones. One type of environmental variability is frequency dependent selection: where a phenotype’s fitness depends on the composition of other phenotypes in the population. Situations where an animal’s fitness is influenced by the outcomes of interactions with other animals exemplify frequency dependent selection, and these situations are often modeled using techniques in game theory. In another paper (Smead and Zollman, 2008), we model the effect of natural selection on plastic individuals using game theoretic techniques.

Because of its axiomatic approach, our earlier model considered games and phenotypes at a coarse level of detail. We here take the other tack by considering one particular game and plastic phenotype. This allows us to investigate different aspects of the Baldwin effect in detail.¹ The game we analyze, the Lewis signaling game, has been used extensively to model the initial emergence of language (Barrett, 2008; Huttegger, 2007; Skyrms, 1996, 2008; Zollman, 2005). This coincides with one of the most common situations where the Baldwin effect is thought to have occurred (e.g. Deacon, 1997;

¹This investigation is also useful because the learning rule we discuss here falls outside the scope of those considered in (Smead and Zollman, 2008).

Dennett, 1991, 1995; Pinker, 2000) and where it is most plausible (Godfrey-Smith, 2003).

Along these lines, Dor and Jablonka (2000) offer a framework for explaining the evolution of language. They describe a complicated and dynamic interaction between learning, genetic assimilation, changing environments, and culture. This framework involves, among other things, “culturally-driven genetic assimilation” by means of an underlying Baldwin effect. Exploring the Baldwin effect in the context of the simple Lewis signaling game will allow us to form an initial assessment of the possible role(s) of the Baldwin effect in the evolution of signaling and language. Along the way, we find that this simple model allows us to separate three distinct aspects of the Baldwin effect.

The first of these distinct aspects is the *Simpson-Baldwin effect*, named for Simpson (1953) by Ancel (1999). The Simpson-Baldwin effect focuses on the effect of evolution on plasticity and predicts that, under certain conditions, initially acquired traits will become genetically assimilated as plasticity is first selected for and then later selected against.² The second aspect, the *Baldwin expediting effect* (also named by Ancel) focuses on the effect of plasticity on the speed of evolution. Here it is predicted that the presence of plastic individuals will result in the emergence of optimal behavior sooner than would be expected without plasticity. Discussion of both of these two aspects of the effect often have focused on single peaked fitness landscapes where there is a unique globally optimal phenotype and no other locally optimal ones. When the possibility of locally-optimal, but globally sub-optimal phenotypes is introduced, another aspect of the Baldwin effect comes into focus. Mills and Watson (2006) demonstrate that the Baldwin effect can allow evolution to “cross fitness valleys” and that the canalization of the Simpson-Baldwin effect is not required to do this. This suggests that the presence of plastic individuals alters the trajectory of evolution by directing the population away from sub-optimal equilibria and toward the global optimum. This last aspect we name the *Baldwin optimizing effect*.

We find that in the model of language we consider, both the Simpson-Baldwin and Baldwin optimizing effect occur, but the Baldwin expediting does not. This separation of the different aspects coincides with a very different set of models (Ancel, 1999, 2000) in which the Simpson-Baldwin

²This aspect has also been explored by Mayley (1996) and genetic assimilation of acquired characters is the focus of Waddington (1953).

effect is found more plausible than the Baldwin expediting effect.

1 The model

While human language has many complicated components, one that has received much recent attention is the arbitrary assignment of signs to objects. This conventionality of language presents an interesting evolutionary circumstance, since there are multiple solutions which appear equally good and many other solutions which are less than optimal. This central feature of animal (and human) signaling is captured in a game described first by David Lewis (1969).

1.1 The Lewis signaling game

In the Lewis signaling game one person (the sender) is given some information about the world (the *state*). He has at his disposal a set of signals that he can send to another player: the receiver. The receiver observes the signal, but not the state, and chooses from a set of actions. There is a single action which is appropriate for a given state. If the receiver takes the appropriate action, both the sender and receiver obtain a reward, otherwise they receive nothing.

In such games it is obviously in both the sender's and receiver's interest to coordinate on a conventional association between states, signals, and acts. In games where there are an equal number of each, any one-to-one map of states to signals and signals to acts which results in the correct act in each state is a Nash equilibrium of the underlying game.

These optimal strategies, dubbed *signaling systems*, represent one potential end point for the process of evolution. However, there are other Nash equilibria that are less than optimal. For instance, if the sender always sends the same signal in every state, the receiver's best response is to take the action which is most often correct regardless of signal. If the receiver is doing this, all potential sender strategies are equally good. This represents another possible Nash equilibrium known as a *pooling equilibrium*.

When there are more than two states, signals, and acts, there are also equilibria which lie in between total pooling equilibria and signaling systems. So-called *partial pooling equilibria* involve the successful communication of

some states, and the pooling of others. These equilibria too are potential end points of the evolutionary process.

Much recent work has been done to determine when these suboptimal states will arise in models of evolution and learning. It is clear that they will sometimes arise and that they represent a potential roadblock to the successful evolution of optimal languages (see Huttegger and Zollman 2009 for an overview).

1.2 Replicator dynamics

One of the simplest models of the evolutionary process is the replicator dynamics (Taylor and Jonker, 1978). This set of differential equations has been used extensively in game theory to model the evolution of strategies under natural selection or similar circumstances and has been used to analyze the evolution of populations playing the Lewis signaling game (Skyrms, 1996; Huttegger, 2007; Huttegger et al., 2007; Pawlowitsch, 2008).

The continuous time replicator dynamics captures the idea that a strategy which does better than other alternatives will grow while those that do worse shrink.³ The dynamics assumes that the population is effectively infinite and that individuals interact with others in the population at random.

In the Lewis signaling game it has been shown that from a random initial starting point, the replicator dynamics can converge to sub-optimal signaling. In fact, only in the case of two states, two signals, and two acts, where the states are equiprobable, is optimal signaling guaranteed to evolve (Huttegger, 2007). All other cases involve some possibility (perhaps small) that populations will become stuck in sub-optimal equilibria (Huttegger, 2007; Huttegger et al., 2007; Pawlowitsch, 2008).

1.3 Herrnstein reinforcement learning

The models using the replicator dynamics suppose a collection of phenotypes which represent strategies in the signaling game. These strategies are not adaptive; they do not change in response to the strategy they are interacting with nor do they change in response to states of the population. However, some recent attention has been given to a particular adaptive strategy, Herrnstein reinforcement learning.

³The replicator dynamics is just one of a class of dynamics that captures this idea. The class is known as payoff-monotone dynamics.

Implementing Herrnstein’s (1970) matching law, Herrnstein reinforcement learning postulates that the probability an individual acts in a certain way is proportional to the past rewards for that action. At each stage an individual has weights representing the past rewards for a given action. An individual determines the probability for each action by dividing the rewards from that action by the total rewards for all actions. The individual then chooses an action at random using these probabilities.

It has recently been proven that, like replicator dynamics, Herrnstein reinforcement learning will converge to optimal behavior in two state, two signal, two act signaling games with equiprobable states (Argiento et al., 2009). Simulation results suggest, however, that it can converge to sub-optimal (partial) pooling equilibria in other circumstances (Barrett, 2006; Skyrms, 2008).⁴

These models have primarily considered two individuals repeatedly interacting with one another, where both individuals are updating behavior via Herrnstein reinforcement learning. They do not consider populations of players, nor the possibility that some individuals might be implementing fixed strategies.

Our model will do just this. We will begin with a population composed of individuals who play one of the available pure strategies (contingency plans in the signaling game) and one type that implements Herrnstein reinforcement learning. Each generation, individuals will be paired to play the game against one another for one thousand iterations of the game. During this time the pure strategy types will unreflectively implement their strategy, and the Herrnstein reinforcement type will attempt to adapt to her opponent. After the iterations of game play, the population will evolve according to a discrete-time version of the replicator dynamic with respect to their type (either using a pure strategy or Herrnstein reinforcement learning).⁵ The dynamics dictate that if the frequency of reinforcement learners in the population will change

⁴These models differ slightly from ours in that they consider so-called “act-based” learning where individuals treat each state or signal as a learning situation distinct from other states or signals. In that way individuals are not learning total contingency plans for every possible state or every possible signal, but only particular moves. Our use of Herrnstein reinforcement presumes that individuals are learning on entire contingency plans.

⁵The discrete time replicator dynamics is more conducive to simulations than the related continuous time dynamics. It should be noted, however, that there are some differences between the two dynamics which may be significant.

proportionally to the ratio of their performance to the population average. Likewise for each non-plastic strategy.

This model will provide us an opportunity to consider the various aspects of the Baldwin effect using a concrete model of language evolution. We will begin with the Simpson-Baldwin effect before turning to the Baldwin expediting and optimizing effects.

2 The Simpson-Baldwin effect

The Simpson-Baldwin effect focuses primarily on natural selection's influence on the evolution of plastic behavior. This version of the effect claims that one should expect, after environmental change, that plastic individuals will first arise and then later be reduced in a population as an acquired trait becomes genetically assimilated. When the environment changes in a relevant way, plastic individuals who are capable of responding to the new environment in better ways can out-compete non-plastic individuals adapted to the previous environment. As a result, plastic individuals are superior to the ancestral types and will be selected for. However, if a type arises later who deterministically engages in the appropriate behavior it will likely be superior, and thus out-compete the (now ancestral) plastic types. By means of a process like this, a trait that is initially learned or acquired in a population becomes genetically canalized.

There are a number of concerns about this story; we will discuss two here. The first focuses on the comparison between plastic types and the behaviorally equivalent non-plastic types. Why should it be the case that non-plastic types are *ceteris paribus* superior to plastic types? It is often supposed that plasticity carries with it a cost, either an exogenous cost of maintaining or developing the mechanism required for plasticity or an endogenous cost generated by delay or failure to implement the optimal strategy. Certainly this is the case with humans; our plasticity comes at great cost.

The second, and perhaps more serious concern focuses on the order of invasions. Godfrey-Smith (2003) points out that if the deterministic behavior occurs before the plastic behavior, it will invade and one will see no growth of plasticity. Why should it be, he asks, that we ought to expect to find plasticity arising first? After considering a few answers, he suggests that the only plausible situation is one where plasticity *creates* the circumstance where the later deterministic types are superior. He cites the evolution of

language as an example. In the ancestral population no one is communicating, thus those who have a genetic predilection for language will do no better. Plastic individuals capable of acquiring linguistic abilities, however, maybe able to arise, learn to communicate with each other, and then take over the population. Once they have done so, their presence makes those with canalized linguistic abilities superior, and thus a population of plastic individuals can be invaded.

While the evolution of language represents one circumstance where the presence of plastic individuals can change the fitness of other types, it is only one example of a more general phenomenon known as frequency dependent selection. Godfrey-Smith’s suggestion applies equally well to all these cases, and so we might be interested in inquiring about frequency dependent selection. We will begin by summarizing the results of our earlier paper (Smead and Zollman, 2008) which reveal Simpson-Baldwin-like phenomena, and then turn to Lewis signaling games (as a model of language). In all these cases, it appears that such effects may be common.

2.1 Plasticity and Evolution

In another paper (Smead and Zollman, 2008), we model the effect of selection on behavioral plasticity (or learning) in the context of social interaction. In these models, a large population of agents are randomly paired to play a repeated game with one another. The agents are either hard-wired to play a fixed strategy of the game or use an *adaptive strategy*. An agent using the adaptive strategy learns to play a best response to the behavior of her co-player in the repeated game. The specific method by which these individuals learn is not specified in these models, only the outcome or aim of the learning process. Two different ways of imposing costs are considered. The first fits with the energetic or developmental aspects of learning where the costs are outside of the interactions (exogenous cost). The other captures the idea that costs arise as a result of errors in the learning process (endogenous cost).

When costs are imposed only on the learners exogenously, the prospects for evolving adaptive individuals in this context are dim. Learning is rarely evolutionarily stable, and only in a restricted subset of games can learners co-exist with non-learners in stable populations. In Lewis signaling games, these results entail that when there is exogenous cost (no matter how small), learning will always be eliminated by evolution.

This demonstrates the second half of the Simpson-Baldwin effect, that

	<i>Cooperate</i>	<i>Defect</i>
<i>Cooperate</i>	2, 2	0, 3
<i>Defect</i>	3, 0	1, 1

Figure 1: Prisoner's Dilemma

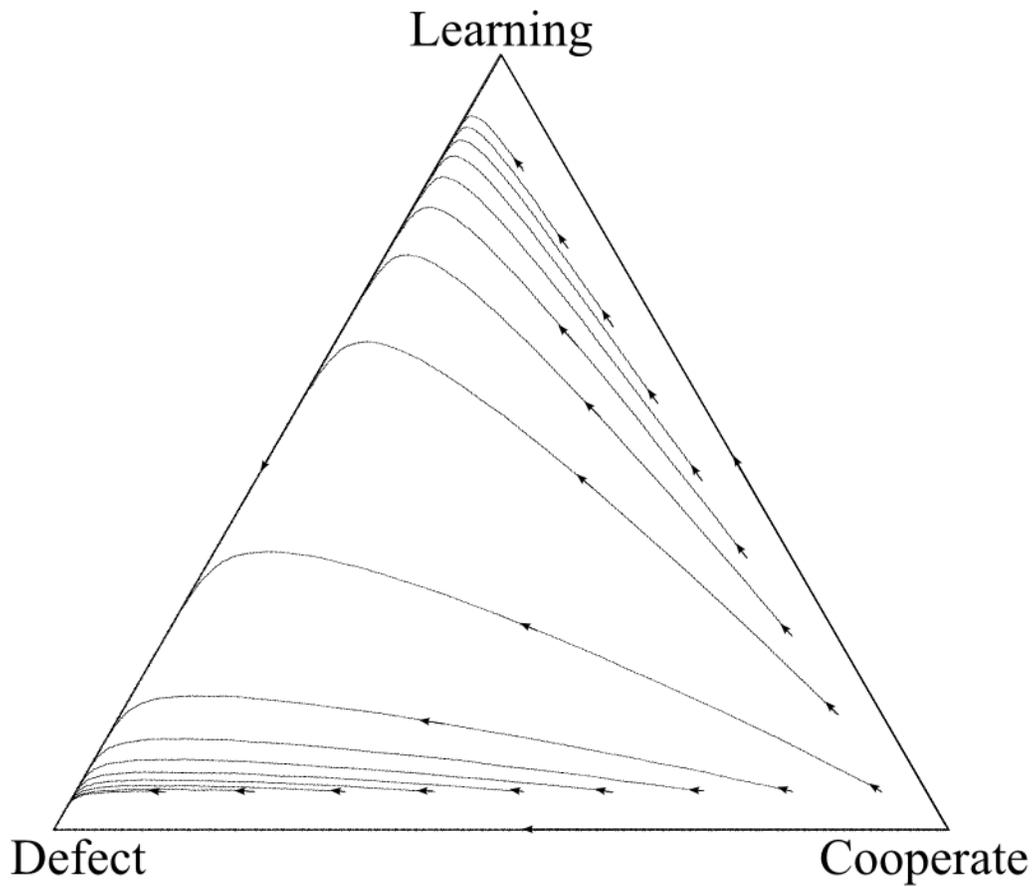


Figure 2: The Prisoner's Dilemma with Plasticity

plasticity is invaded by a canalized type. We also find the first component, that plasticity invades other populations. Consider the prisoner’s dilemma (see figure 1). In this game there are two canalized types, cooperators and defectors. We also add a plastic type: learners. Figure 2 shows the replicator dynamics this situation. The evolutionary paths that begin in populations largely composed of cooperators are initially invaded by learners (who learn to defect in the game) before the learners are eliminated in favor of non-learning defectors.

The connection between the Simpson-Baldwin effect and this result must be explained in more detail. The Simpson-Baldwin effect is usually described as the genetic canalization of *novel* traits and the trait that is canalized in this model was already present in some non-plastic individuals. Modeling genuine novelty is difficult in the framework of evolutionary game theory. However, there is a sense in which the phenomena seen these models are a generic version of the Simpson-Baldwin effect, where the emergence of novel traits is a special case. The Simpson-Baldwin effect with respect to novelty can be understood as a population which is evolving on or near the edge of the simplex. Using the Prisoner’s Dilemma example, we can imagine a population composed of all *Cooperate* and a few plastic individuals. The plastic individuals will learn to play *Defect*, which is a novel behavior for this particular population. Once learners take over the population, a mutation the direction of *Defect* will cause the population to proceed to be taken over by individuals who are non-plastic and always play *Defect*.⁶ Thus, the Simpson-Baldwin effect with respect to a novel behavior is a special case of the phenomena described in our previous work.⁷

What if there are no exogenous costs and the only costs come from errors in learning? In this case, we modeled a “cost” as a deviation from the best response by the learners.⁸

In a highly restricted setting (focusing only on 2×2 games), we find

⁶Here, it is not specified *how* the trait becomes canalized. This may be important for determining the plausibility of some particular evolutionary path, but is beyond the scope of this paper.

⁷The fact that there are only two possible behaviors in this particular examples can make the interpretation of “novelty” seem artificial. However, similar interpretations and examples could be given for much larger games with many more possible actions.

⁸In this case, a “cost” could actually be beneficial, if deviations somehow resulted in higher payoffs accidentally. The cost is quantified here by looking at the proportion of plays that deviate from the “normal” best-response behavior of the learners and calculating the payoff difference between the normal behavior and the cases where deviation occurs.

that it is unlikely for plastic individuals to evolve. To make analytic results tractable in this case, it was assumed that the probability of deviation from the normal behavior is fixed regardless of opponent. Neither this assumption nor the restriction to 2×2 games applies to Herrnstein reinforcement learning in signaling games. A signaling game cannot be represented as a 2×2 game, and in the short run, the rate of exploration or deviation by Herrnstein reinforcement learners depends heavily on the behavior of the co-player. In some games, this means that a population of reinforcement learners can resist invasion from pure strategy players.

To highlight this second point, consider a population of individuals paired to repeatedly play the Prisoner’s Dilemma. The types of individuals are cooperators, defectors and Herrnstein reinforcement learners. Consider the behavior of learners in the short run against each of the three types. Eventually, learners will be driven to play *Defect* regardless of opponent, since it is strictly dominant. But in the short run learners will play *Cooperate* with one another and achieve a higher payoff than *Defect* against *Defect*.⁹ With only endogenous costs (conceived as generated by exploration), reinforcement learning can be evolutionarily stable. Figure 3 shows the evolutionary dynamics in this case.

With endogenous costs to learners, there are settings (such as the Prisoner’s Dilemma) where we should not expect a reduction in strategic plasticity, but rather that Herrnstein reinforcement learning can evolve and dominate the population. This consideration leaves open the possibility that we may not see a Simpson-Baldwin effect in signaling with respect to reinforcement learners.

In a very different model Suzuki and Arita (2004) consider a repeated Prisoner’s Dilemma which includes plastic strategies. They also observe a Simpson-Baldwin effect, where plasticity is selected for and then later reduced. In their model, however, plasticity is not eliminated and most individuals are partially plastic. This later possibility is not included in our models, individuals are either plastic or not and so as a result represents a rather different way of modeling plasticity.

⁹We should expect Herrnstein reinforcement learners to take longer to play *Defect* regularly against cooperators than against defectors, and likewise longer against other reinforcement learners than defectors. Hence, the “error rate” of reinforcement learners is dependent on the strategy being used by the co-player, and thus this example is not subject to the theorems regarding endogenous cost in (Smead and Zollman, 2008).

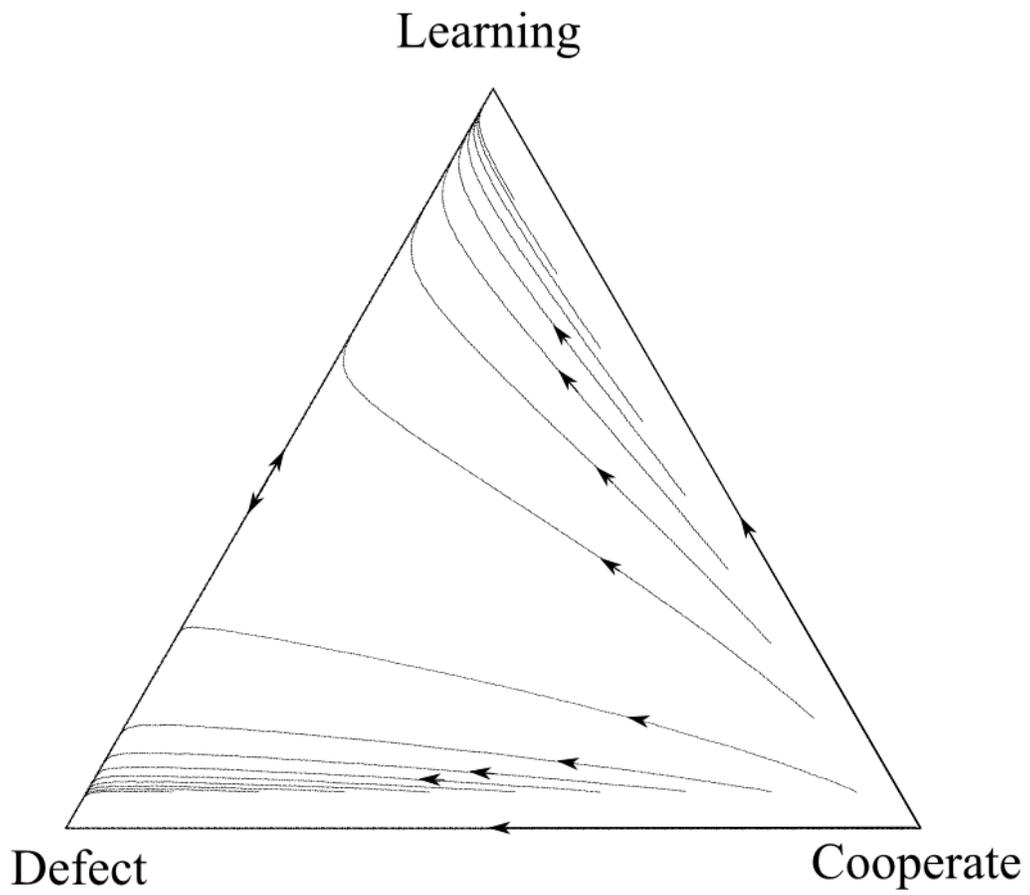


Figure 3: The Prisoner's Dilemma with Plasticity

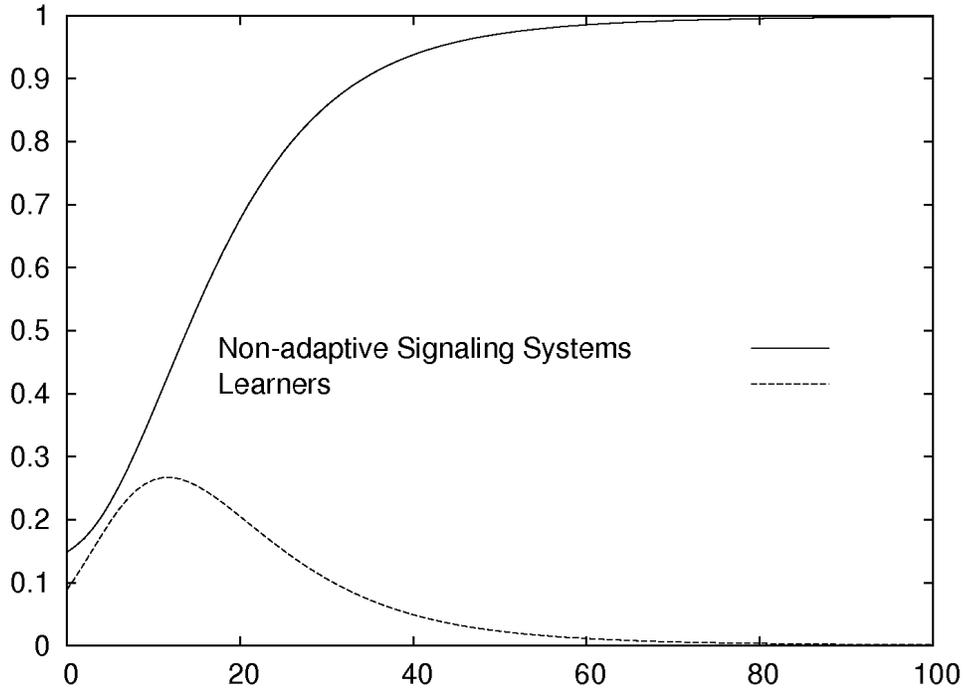


Figure 4: Simpson effect

2.2 Signaling games

Here we consider Herrnstein reinforcement learning as a type in a population of other types playing a signaling game. Two individuals are paired to repeatedly play a signaling game with one another and they will receive the average payoff for their interaction.¹⁰ There is no exogenous cost to learning, but Herrnstein reinforcement learning will make errors. Since this is a partnership game (both player receive the same payoff) errors are costly—learners cannot gain from them.

Figure 4 shows the average proportion of learners over time for 10,000 separate starting points of a two state, two signal, two act signaling game with equiprobable states. Here we see a distinct Simpson-Baldwin effect. Initially the Herrnstein reinforcement learners do well, increasing in the population.

¹⁰This average payoff was determined by running 10,000 independent runs of Herrnstein reinforcement against each pure type and constructing an expanded game using these estimated payoffs.

Learners do well against all types, they learn to signaling effectively with those that signal and they do as well as they can against those that do not. However, once there is a significant number signalers (both learners and pure strategy signalers) in the population, those that deterministically play one signaling system will eventually take over. The reason is that, when interacting with a learner, it is better for an individual to have a fixed signaling strategy and allow the learner to adapt than it is for both individuals to try and adapt simultaneously. Two learners take longer to find successful signaling than one learner and one pure-strategy signaler.

At a behavioral level, a population with a significant proportion of learners will acquire the ability to communicate by learning. Once pure strategy signaling types begin to invade, the ability to effectively communicate becomes innate. The end result of the evolutionary process is a population that uses a single signaling system without aid of learning. Which signaling system invades is a matter of chance and it is determined by the initial proportions which are determined at random. But eventually one or the other wins out.

Our caution about novelty occurs here again. There are signaling system types present in most of these populations, which means that the signalers are not using a novel strategy. But, as in the case of the Prisoner’s Dilemma, we find a more general Simpson-Baldwin effect where learners are selected for and then selected against even in populations with both deterministic and plastic types.

3 Baldwin expediting and optimizing effect

Another aspect of the Baldwin effect focuses on how the presence of plastic phenotypes may alter the course of evolution by “speeding up” or “expediting” the evolutionary process. Here it is conjectured that the presence of plastic individuals will result in the attainment of optimal phenotypes more quickly than would have occurred without the presence of plastic types in the population (the *expediting effect*). More radically, some have suggested that the presence of plastic individuals helps populations to find optimal behaviors that would not have been found without their presence (the *optimizing effect*).

Most studies have focused on the expediting rather than the more-radical optimizing effect. Mills and Watson (2006) provide one discussion on the

ability of the Baldwin effect to cross fitness valleys. Daniel Dennett also seems to come close to discussion of the optimizing effect, but he appears to vacillate between the two.¹¹ Other examples discussed in the literature tend to focus on fitness landscapes with single peaks, that is situations where there is a unique globally optimal behavior and no locally optimal behaviors that are not globally optimal (Ancel, 2000; Dennett, 1991, 1995; Hinton and Nowlan, 1987).¹²

But, this is not the case in generalized signaling games. Most signaling games have stable local optima which are not globally optimal. So, we must be careful to distinguish these effects. The Baldwin expediting effect (in isolation) would require that the ultimate basins of attraction of different end states would remain the same irrespective of the presence of a learning phenotype, but the system would converge to these states more quickly with learning than without. The Baldwin optimizing effect on the other hand would entail that the size of the basins of attraction would be altered by the presence of learning as well.

First, we will consider the Baldwin expediting effect. To remaining consistent with the previous literature on this subject we will consider a two state, two signal, two act signaling game with equiprobable states. This game has two global optima and no stable local optima (Huttegger, 2007). Figure 5 shows the results of simulations comparing the evolution of populations with and without learning respectively. The y-axis represents the average payoff of the population. While the population with learners starts out closer to optimal behavior (by mere fact that there are more types who behave optimally) we see that the slope of the line for learners is more shallow which represents slower evolution to the optimum.

This suggests that, in the context of signaling games, the Baldwin expediting effect is not present. Learners do not assist the population in converging to optimal behavior more quickly. Here, however, concerns about novelty again appear. Those who describe the Baldwin expediting effect suggest that the presence of plastic individuals help to find phenotypes which are not

¹¹For instance, in his (1991) he says, “If it weren’t for the plasticity, however, the effect wouldn’t be there” (186). But in his (1995) he weakens this claim, “...their species will *evolve faster* because of its greater capacity to discover design improvements in the neighborhood” (79, emphasis in original).

¹²Suzuki and Arita (2004) are an exception. They consider strategies in a repeated Prisoner’s Dilemma which include plasticity. Although they do not single it out, they observe a Baldwin optimizing effect in their model as well.

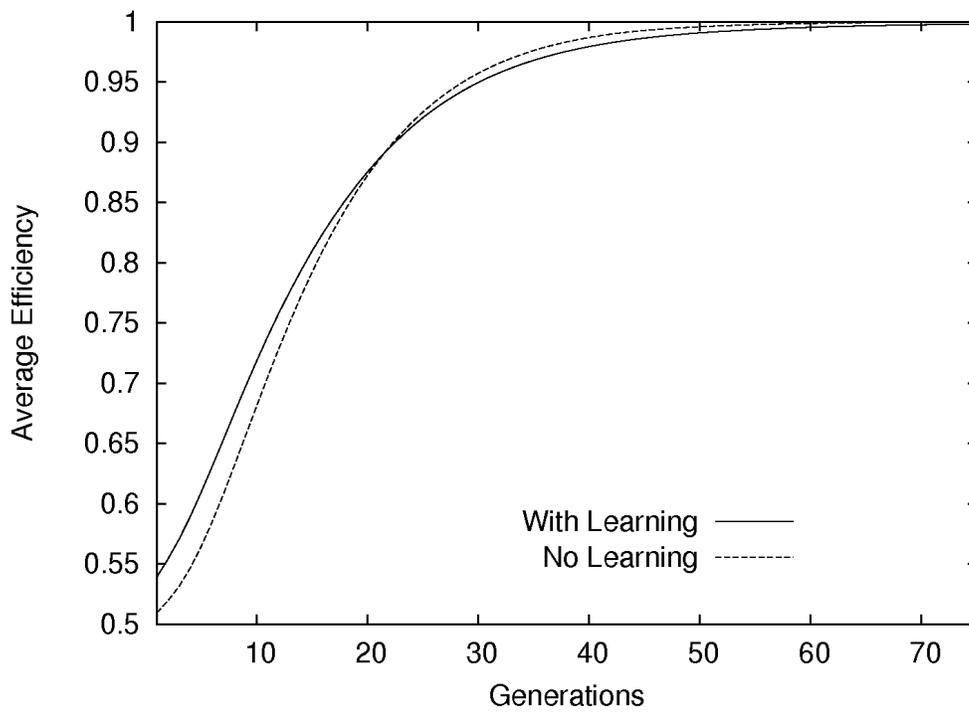


Figure 5: Expediting effect

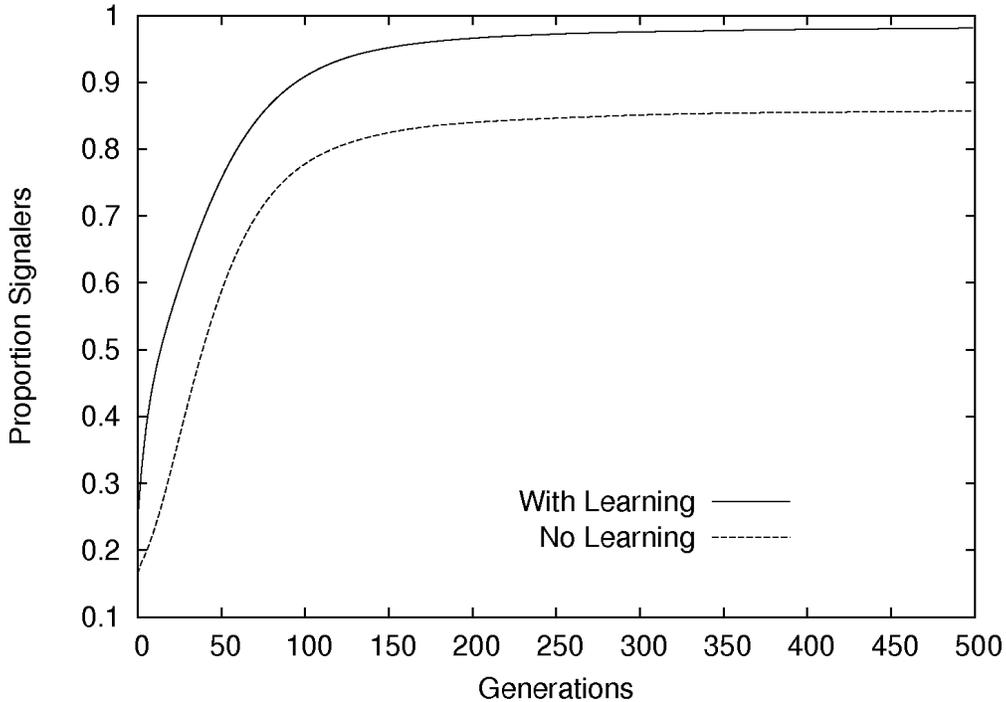


Figure 6: Optimizing Effect

present in the ancestral population. In these simulations, we choose an initial population state at random which creates a population with some optimal types, so no novel behavior is generated. As a result, the negative result here should not be generalized to conclusions about the Baldwin expediting effect elsewhere.¹³

The Baldwin optimizing effect does occur in signaling games with sub-optimal equilibria. Figure 6 shows an example of a two state, two signal, two act signaling game with non-equiprobable states (the probabilities are 0.8 and 0.2 respectively). This game has local optima which are not globally optimal as described in section 1.1. The y-axis in this figure represents

¹³Our informal analysis of novelty presented above does not answer the question here, since we would need to measure directly the convergence speed in different models. This is beyond the ability of the replicator or replicator-mutator dynamics, and would instead require a finite population model. Models such as the Moran process (Moran, 1962) might be appropriate for this task. Such a model would conform to our informal discussion of novelty, but space prevents us from discussing it here.

the average proportion of the population which contains optimal strategies. Optimal strategies are present more often after a few generations and pooling equilibria are avoided. The learners are eventually eliminated, but their early presence has altered the evolutionary trajectories of populations that would have otherwise evolved toward the pooling equilibria.

Perhaps the most surprising feature of this result is that learners themselves are not guaranteed to converge to optimal behavior. Two individuals both employing Herrnstein reinforcement learning in this game will often converge to the sub-optimal total pooling equilibria. However, in the limit, a Herrnstein reinforcement learner playing against a fixed signaling system strategy will learn the appropriate signaling system. This latter feature appears to be sufficient to move populations inside the basin of attraction of signaling systems before the learners are eliminated.

4 Conclusion

Several previous mathematical and simulation models of the Baldwin effect have provided a variety of answers regarding the plausibility of the effect in different circumstances (Ancel, 1999, 2000; Hinton and Nowlan, 1987). These models, along with much of the discussion of the Baldwin effect, have primarily focused on a changing external environment, on adaptive problems which only require adaptation to the external environment, and on a fitness landscape without sub-optimal equilibria. We have analyzed a circumstance that differs in these three respects. Our model has a stable game, but the fitness of different types changes according to changes in the population; there is no strategy which is optimal regardless of the strategies used by others in the population. In addition, there are equilibria which are not globally optimal.

We have focused on a model of the evolution of language already well studied by philosophers. Our aim was to evaluate the plausibility of the different aspects of the Baldwin effect in this domain. Ultimately we conclude that some aspects are very likely.

The Simpson-Baldwin effect appears to be present in many models of frequency dependent selection, including our model of the evolution of language. While many have focused on the emergence of novel traits via plasticity, our model shows that the Simpson-Baldwin effect may occur even in situations where the trait being implemented by plastic individuals is not novel. In the

early stages of evolution, effective signaling is often generated by learning and as evolution continues, learners are driven out in favor of individuals that effectively signal without learning. The growth and later reduction of plasticity by frequency dependent selection may be a very widespread phenomena.

We find little support for the Baldwin expediting effect, but because our model cannot account for novelty in the way discussed by many authors, this negative conclusion is not particularly telling. Instead, we are left with little to say about this aspect of the Baldwin effect.

Finally, we distinguished a third effect, the Baldwin optimizing effect and found significant support for that effect in our model of the evolution of language. Whether this effect is present in other games is an open question, but its presence in even one is surprising, and it confirms a more general moral – the presence of plastic individuals can effect the trajectory of evolution, even if those plastic individuals are later eliminated.

References

- Ancel, L. W. (1999). A quantitative model of the Simpson-Baldwin effect. *Journal of Theoretical Biology* 196, 197–209.
- Ancel, L. W. (2000). Undermining the Baldwin expediting effect: Does phenotypic plasticity accelerate evolution? *Theoretical Population Biology* 58, 307–319.
- Argiento, R., R. Pemantle, B. Skyrms, and S. Volkov (2009, February). Learning to signal: Analysis of a micro-level reinforcement model. *Stochastic Processes and their Applications* 119(2), 373–390.
- Barrett, J. A. (2006). Numerical simulations of the Lewis signaling game: Learning strategies, pooling equilibria, and the evolution of grammar. Technical Report MBS 06-09, University of California, Irvine: Institute for Mathematical Behavioral Sciences.
- Barrett, J. A. (2008). Dynamic partitioning and the conventionality of kinds. *Philosophy of Science* 74, 527–546.
- Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. Norton.

- Dennett, D. (1991). *Consciousness Explained*. Little, Brown, and Company.
- Dennett, D. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon and Shuster.
- Dor, D. and E. Jablonka (2000). From cultural selection to genetic selection: A framework for the evolution of language. *Selection* 1(1-3), 33–56.
- Godfrey-Smith, P. (2003). Between baldwin skepticism and baldwin boosterism. In B. H. Weber and D. J. Depew (Eds.), *Evolution and Learning: The Baldwin Effect Reconsidered*. MIT Press.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior* 15, 245–266.
- Hinton, G. E. and S. J. Nowlan (1987). How learning can guide evolution. *Complex Systems* 1, 495–502.
- Huttegger, S. (2007, January). Evolution and explanation of meaning. *Philosophy of Science* 74(1), 1–27.
- Huttegger, S., B. Skyrms, R. Smead, and K. Zollman (2007). Evolutionary dynamics of lewis signaling games: Signaling systems vs. partial pooling. *Forthcoming in Synthese*.
- Huttegger, S. and K. J. Zollman (2009). Signaling games: The dynamics of evolution and learning. In A. Benz (Ed.), *Language, Games, and Evolution*.
- Lewis, D. (1969). *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- Mayley, G. (1996). Landscapes, learning costs, and genetic assimilation. *Evolutionary Computation* 4(213-234).
- Mills, R. and R. Watson (2006). On crossing fitness valleys with the Baldwin effect. In *Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems*, pp. 493–499.
- Moran, P. (1962). *The Stastical Process of Evolutionary Theory*. Oxford: Clarendon Press.

- Pawlowitsch, C. (2008). Why evolution does not always lead to an optimal signaling system. *Games and Economic Behavior* (63), 203–226.
- Pinker, S. (2000). *The Language Instinct*. HarperCollins.
- Simpson, G. G. (1953, June). The Baldwin effect. *Evolution* 7(2), 110–117.
- Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge: Cambridge University Press.
- Skyrms, B. (2008). *Signals: Evolution, Learning, and the Flow of Information (book manuscript)*.
- Smead, R. and K. Zollman (2008). The stability of strategic plasticity. *Carnegie Mellon University Department of Philosophy Technical Report CMU-PHIL-182* .
- Suzuki, R. and T. Arita (2004). Interactions between learning and evolution: The outstanding strategy generated by the Baldwin effect. *BioSystems* 77, 57–71.
- Taylor, P. and L. Jonker (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences* 40, 145–156.
- Waddington, C. H. (1953, June). Genetic assimilation of an acquired character. *Evolution* 7(2), 213–234.
- Zollman, K. J. (2005). Talking to neighbors: The evolution of regional meaning. *Philosophy of Science* 72, 69–85.