

Social Interaction and the Invasion of Adaptive Strategies

Rory Smead

May 8, 2009

Abstract

This paper uses the methods of evolutionary game theory to investigate the effect of social interaction on the evolution of learning very broadly construed as strategic plasticity (or “adaptive strategies”). We delineate the conditions under which social interaction alone, apart from any additional external environmental variation, can provide the selective pressure necessary for the evolution of costly plasticity. Furthermore, it is argued that we should not expect traditional learners that “best-respond” to dominate the population and that it may be important to consider non-traditional learners in studying the evolution of learning in these settings.

1 Introduction

Cognitive systems are generally biologically expensive, delicate, and complex. How did such systems evolve? The primary function of cognitive systems is to enable adaptation; so one would assume that complex environments will provide the best setting for their evolution: this is the *environmental complexity thesis* advanced by Peter Godfrey-Smith (Godfrey-Smith, 1996, 2002a,b). Indeed, there has been much interesting work on the evolution of learning in complex environments.¹ But, there are many forms of environmental complexity. In addition to ecological change and a widely varied

¹This work includes Borenstein and Ruppin (2006); Borenstein et al. (2008) and a number of studies related to the Baldwin effect in biology (Baldwin, 1896; Ancel, 1999, 2000).

environment, Godfrey-Smith includes interaction with other agents as a form of environmental complexity.

This latter form of environmental complexity, social interaction, is in some ways much different than the others mentioned and some scholars argue that social intelligence or *Machiavellian Intelligence* may be the key to understanding the evolution of cognition (Humphrey, 1976; Byrne, 1997; Byrne and Bates, 2007). The idea is that intelligence evolved to solve social problems. Others that have studied the evolution of learning in social interaction studies include Harley (1981, 1983, 1987); Maynard Smith (1982); Rogers (1988); Kirchkamp (1996, 1999); Richerson and Boyd (2000); Rodríguez-Sickert et al. (2008). Imitative learning in particular has caught the attention of many scholars including Schlag (1998); Sterelny (2003); Borenstein and Ruppin (2006). While these studies are interesting, the current paper will take a step back relative to the settings typically considered and focus on the conditions for the initial invasion of adaptive individuals with respect to social interaction.

Questions around the evolution of (human) cognition have obvious appeal. However, intelligence is just one form of plasticity that occurs in nature, and similar questions will arise regarding the evolution of plasticity generally. What role does social interaction or social complexity play in the evolution of plasticity, or perhaps more specifically *strategic plasticity*? To answer this question, we will use the tools of evolutionary game theory and strategic plasticity will be represented as a learning strategy or an *adaptive strategy*. An adaptive strategy broadly conceived provides an agent some mechanism of modifying her behavior over time.²

The focus of this paper will be on an idealized case: determining if and when social interaction *alone* can allow for the invasion and evolution of strategic plasticity. Despite the idealization, this issue is a complex one: the best way for one agent to learn or adapt will depend on the way that other agents are learning. As we will see, the ways in which social interaction can provide selective pressure for the evolution of learning is nuanced and complex. Nonetheless, in some general settings, a precise characterization the conditions for the evolution of learning in strategic settings is possible.

²We will use the terms “learning”, “adaptive strategy”, and “plasticity” interchangeably: the models provided below are intended to be applicable to any variety of social plasticity. Whether or not the term “learning” is appropriate in a particular situation will most likely depend on the way in which plasticity is achieved, which will not be specified in the models considered below.

We provide a step toward a general theoretical analysis of the evolution of learning in settings of social interaction and consider the implications for the study of learning, plasticity and social evolution. After briefly presenting some preliminaries, the subsequent analysis is broken into two parts based on how we characterize the learning situation: learners adapting to a population and learners adapting to other individuals.

We then identify the conditions required for social interaction alone to allow for the invasion of learners (flexible individuals) relative to inflexible individuals. Situations that involve unstable or polymorphic populations are most conducive to the evolution of learners and allowing learners to respond to other individuals also creates conditions where learners can thrive in social interactions. Furthermore, we will argue that there may be good reason to expect settings of social interaction select for “non-traditional” (non-best-response) learning strategies or populations that are polymorphic with respect to learning. Overall, these results reveal that social interaction may be an important force in the evolution of learning as well as raise methodological questions with respect to modeling social evolution.

1.1 Preliminaries: Evolutionary Game Theory

In the models presented here, social interaction will take the form of some 2-player symmetric game (labeled G), which consists of a set of strategies S for each player and a payoff function $\pi : S \times S \rightarrow \mathbb{R}$ which specifies the payoffs for every pair of strategies. The set of mixed strategies will be represented as ΔS and payoffs with respect to mixed $\sigma \in \Delta S$ will be calculated as expected utilities. Let $Br : \Delta S \rightarrow \Delta S$ be a correspondence such that $Br(\sigma) = \sigma'$ where $u(\sigma', \sigma) \geq u(\sigma'', \sigma)$ for all $\sigma'' \in \Delta S$ (Br is the best-response correspondence). A strategy profile (σ, σ') is a *Nash equilibrium* just in case each strategy is a best response to the other.

We will be interested in a specific class of individuals: learners. Since “learning” carries many meanings, we will focus on a specific meaning with respect to social interaction. A learner will be an individual that employs some *adaptive strategy* (represented by \mathcal{L}_i), which is capable of using different strategies at different times on the basis of some response rule. So as to ensure general results, we will say very little about the character of the response rules of adaptive strategies unless we are considering specific examples.³ In each

³We will also say very little about the sort of input/information available to the learners.

model, we will assume an infinite population of randomly-mixing individuals that use either pure strategies or adaptive strategies. We will also assume that each \mathcal{L}_i has an associated cost c_i .⁴ The idea will be to specify the conditions under which adaptability generates benefits that can overcome the associated costs of being plastic.

The adaptive strategies will be introduced to the game G by creating a “learning game” G^L which expands the strategy set to S^L and consists of both adaptive strategies and pure strategies. The payoffs π^L of G^L will be a function of the payoffs of G and the description of the learning situation (described in detail within the models below). The evolutionary analysis will be applied to G^L using the standard concepts of evolutionary stability and related concepts of invasion by mutants. As a notational convention, unlabeled strategy sets, payoff functions, etc. will refer to the base game G and those that refer to G^L will be labeled with an L .

Before continuing on to the models, it will be helpful to outline other important concepts for evolutionary analysis in games.

Definition 1. *A strategy s^* is an evolutionarily stable strategy (ESS) if and only if*

1. $u(s^*, s^*) > u(s, s^*)$ for all $s \neq s^*$ **or**
2. $u(s^*, s^*) = u(s, s^*)$ and $u(s^*, s) > u(s, s)$ for all $s \neq s^*$.

The concept of the ESS captures an “uninvadable” strategy and is due to Maynard Smith and Price (1973).

In some cases it will be helpful to analyze the game on a population-level. We will imagine an infinite population of randomly mixing individuals that each have some strategy of G^L . For a game with n strategies, the space of possible populations is simply:

$$P^n = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n \mid \sum_{i=1}^n x_i = 1\}.$$

Note that a population $x \in P^n$ is just a vector that is identical to some mixed strategy $\sigma \in \Delta S$. Thus we can refer to a strategy’s payoff against a

Instead we will simply specify the range of behaviors available to the learners, which may be determined on a number of factors including what information the learners use.

⁴Where only one adaptive strategy is considered, we will omit the associated index.

population, or one population's payoff against another, in the same way that we can speak of payoffs with respect to mixed strategies. To fit with evolutionary analysis at a population-level, we will use "fitness" rather than utility. We will denote the fitness of strategy (or population) x against population y as follows:

$$F(x, y) = \sum_{j \in S} u(x, j)y_j = \sum_{i \in S} \sum_{j \in S} \pi(i, j)x_i y_j$$

where x_i, y_j represent the proportion of strategy $i, j \in S$ in the support of x, y respectively. A population x will be said to be at a *Nash equilibrium* (NE) just in case all the strategies represented in x are best responses to x . The concept of evolutionary stability can also be applied to populations and is closely related to the definition of ESS. Suppose we introduce some proportion of new individuals into a population x , thus perturbing x by ϵ and creating x' . If x does better relative to the in the perturbed state x' than x' does relative to itself, x is evolutionarily stable:

Definition 2. *A population x is an evolutionarily stable state (ES state) if and only if*

1. $F(x, x) > F(x, x')$ for all $x' \neq x$. **or**
2. $F(x, x) = F(x, x')$ and $F(x, x') > F(x', x')$ for all $x' \neq x$ within a neighborhood of x .

With a single, infinite, and randomly mixing population, ESS and ES states are equivalent.⁵ However, given the character of the two models below it is helpful to set out both concepts.

Since we will be specifically interested in the invasion of strategies, we can use the intuitions behind these stability concepts to introduce the notion of weak and strong invasiveness.

Definition 3. *a) A strategy s is weakly invasive with respect to x (or strategy σ) if and only if $F(s, x) \geq F(x, x)$ and b) s is strongly invasive if $F(s, x) > F(x, x)$.⁶*

⁵For a discussion and analysis on when these two concepts come apart see Bergstrom and Godfrey-Smith (1998) and Thomas (1984).

⁶Note that even if a strategy is weakly invasive it may not be able to invade – this will depend on the second condition for ES state relative to the population. But if a strategy is not weakly invasive, it cannot invade. Weak and strong invasiveness is similar and related to the concept of strong uninvadability introduced by Bomze and Weibull (1995).

The stability concepts above are important, but to get a full evolutionary picture, evolutionary dynamics are also important. One of the most common dynamic used in large, randomly mixing population settings is the replicator dynamic (Taylor and Jonker, 1978):

$$\dot{x}_i = x_i[F(i, x) - \phi].$$

In words, the frequency of type i changes in proportion to the difference between the fitness of type i and the average fitness in the population ϕ . Within standard models, there are many close links between the concepts set out in these preliminaries, for instance, the following set of implications holds with respect to the replicator dynamic in pairwise interactions and linear payoffs: ESS \rightarrow Asymptotic Stability \rightarrow Neutral Stability \rightarrow Lyapunov Stability \rightarrow NE.⁷ Given the complexity of the models considered here, we will forego a full dynamic analysis, and instead consider some illustrative examples that involve simulations with the discrete time version of the replicator dynamic.

Finally, we will be considering the relationships between two games at the population-level with a different number of types: one 2-player symmetric game G and the learning game G^L . It will be convenient to extend a population x that specifies a distribution of strategies in G to the new game G^L in a unique way. To do this, we will simply let a population in G be the corresponding population in G^L where none of the adaptive strategies are present. More precisely, we set $x^L = (x_1^L, \dots, x_{n+m}^L)$ where n is the number of pure strategies in G and m is the number of additional adaptive strategies in G^L . Then set $x_i^L = x_i$ for $1 \leq i \leq n$ and $x_i^L = 0$ for $i > n$. In other words, populations specified for game G are simply populations on one face of the simplex for G^L . Since we will never reduce a population from x^L to x , for notational simplicity we will usually suppress the superscript and refer to a population x while assuming the straight forward extension.

2 Adapting to the Population

There are two extremes to thinking about how agents will be learning to behave in social interaction settings: responding to a population or responding

⁷It should be noted that the presence of learners in the population complicates matters significantly and the usual connections cannot necessarily be assumed. For the details on the relationship between these concepts see Weibull (1997), Hofbauer and Sigmund (1998), and Sandholm (2008).

to individuals. In the former case, agents repeatedly play a game with others in the population but do not distinguish individuals in the population and simply learn how to behave when interacting with others from that population generally. In the later case, agents repeatedly play a game with specific individuals and learn how to interact with other individuals; agents may learn different behavior for different interaction partners. These two points represent the extreme cases of a continuum where agents learn to interact with larger or smaller subsets of a population (or where agents have less or more discriminatory ability). I will examine each extreme and the prospects for the evolution of adaptive strategies in each, beginning with adapting to the population.

In this setting, the learning situation is where agents must learn how to interact with their population as a whole and cannot distinguish between individuals (or types of individuals) in the population. This may be because agents simply cannot distinguish individuals, have no memory of specific interactions, or simply do not have repeated interactions with specific individuals. We imagine learners who have a large number of interactions with others and have an adaptive rule which specifies some behavior given the history of their interactions. We will also make the idealizing assumption that these individuals adapt quickly relative to their life-span so that the payoffs they receive are relatively stable (i.e. the interactions used for learning do not have a large bearing on the overall payoff from being a learner).

To model this situation, we will imagine an infinite, randomly mixing population of individuals who are paired to play a 1-shot game G with one another. This pairing is repeated with new partners many times, each chosen at random independently. There are pure strategy types in the population, which always behave the same way and adaptive individuals \mathcal{L}_i as well, which behave according to some adaptive rule. The evolutionary analysis will occur with respect to a game population G^L which extends the strategy set of G to S^L by adding adaptive strategies $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_n$.

Since we are modeling learners as responding to a population and do not have a simple fixed behavior it becomes necessary to distinguish between the behavior of the population and the distribution of types in the population. If G has m strategies, the space of possible behaviors will simply be:

$$B^m = \{(y_1, \dots, y_m) \in \mathbb{R}_+^m \mid \sum_{i=1}^m x_i = 1\}.$$

Where y_i represents the proportion of times strategy $i \in S$ is played. Without adaptive strategies in the population the behavioral distribution is identical to the population distribution, since strategy types uniquely identify behavior. However, with adaptive strategies this will not be the case. We will let $\beta : P^n \rightarrow B^n$ represent a function that determines a long-run behavioral distribution for a population x according to the distribution of x and the nature of the adaptive strategies present in x .

\mathcal{L}_i will adopt a strategy (possibly mixed) of game G as a function of the distribution of behaviors. Let $r_i : B^m \rightarrow \Delta S$ represent the adaptive function for \mathcal{L}_i such that $r_i(\beta(x)) = \sigma$ means that \mathcal{L}_i adopts the mixed strategy σ of G in response to the behavior $\beta(x)$. This means that the payoffs of G^L will not be fixed but rather change with respect to the behavior and composition of the population, and in possibly very complex ways (since r_i and β are interdependent). This is unusual with respect to the traditional analysis of evolutionary games (in which the payoffs between strategies do not change with respect to the behavior of the population). The fitness functions for G^L will be:

- $F^L(s, x) = F(s, \beta(x))$ for any $s \in S$ and
- $F^L(\mathcal{L}_i, x) = F(r_i(\beta(x)), \beta(x)) - c_i$ for any \mathcal{L}_i .

Where c_i represents the cost (possibly very small) to using adaptive strategy \mathcal{L}_i . This makes the evolutionary analysis of populations containing adaptive strategies impossible without specifying the response rules. However, we are still able to consider whether or not adaptive strategies might be able to invade a population x that has no adaptive strategies in its support. We will refer to the situation described by this model as “adaptive strategies responding to the population.”

We will begin by showing the conditions for an invasion of an adaptive strategy with respect to a learner-less population. The pair of theorems below shows that the key condition for the invasion of adaptive strategies in this setting is the instability of the population.

Theorem 1. *If adaptive strategies respond to the population, $c > 0$ and population x has no adaptive strategies and is at a NE of G then no adaptive strategy \mathcal{L} is weakly invasive with respect to x .*

Proof. Suppose \mathcal{L} responds to the population, $c > 0$ and x is at a NE of G (and that x has no adaptive strategies in its support so $\beta(x) = x$). Without

loss of generality, let $i \in S$ be a strategy in the support of x and $j \in S$ be a strategy the support of $r_{\mathcal{L}}(x)$. By hypothesis, $F(i, x) \geq F(s, x)$ for all $s \in S$, hence $F(i, x) \geq F(j, x)$ for any j in the support of $r_{\mathcal{L}}(x)$. It follows that, $F(i, x) \geq F(r_{\mathcal{L}}(x), x)$. We now have the following:

$$\begin{aligned} F^L(\mathcal{L}, x) < F^L(i, x) & \text{ iff } F(r_{\mathcal{L}}(x), x) - c < F(i, x) \\ & \text{ iff } c > F(r_{\mathcal{L}}(x), x) - F(i, x). \end{aligned}$$

Any $c > 0$ will satisfy the right side of the above equation, and \mathcal{L} will not be weakly invasive to x . \square

This theorem shows that if a population is at a NE with respect to G it cannot be invaded by any adaptive strategy in G^L . This implies that any x which is evolutionarily stable with respect to G cannot be invaded by an adaptive strategy. The next theorem shows that some \mathcal{L} will be able to invade a population that is not at an NE.

Theorem 2. *If adaptive strategies respond to the population, $c > 0$ is sufficiently small, and x has no adaptive strategies and is not at an NE of G then there is some \mathcal{L} that is strongly invasive with respect to x .*

Proof. Suppose \mathcal{L} responds to the population, $c > 0$ and x has no \mathcal{L} (so $\beta(x) = x$) and is not at a NE of G . By hypothesis, there is some s in the support of x that is not a best response to x and hence some $t \in S$ s.t. $F(s, x) < F(t, x)$. Let $\sigma = (y_1, \dots, y_n) \in \Delta S$ be such that $y_i = x_i$ for $i \neq s$ and $y_t = x_s$ (σ is the mixed strategy just like x except it plays t in place of s). By virtue of $F(s, x) < F(t, x)$, $F(x, x) < F(\sigma, x)$. Thus, for any \mathcal{L} such that $r_{\mathcal{L}}(x) = \sigma$ we have the following:

$$\begin{aligned} F^L(\mathcal{L}, x) > F^L(x, x) & \text{ iff } F(r_{\mathcal{L}}(x), x) - c > F(x, x) \\ & \text{ iff } c < F(\sigma, x) - F(x, x). \end{aligned}$$

Since $F(x, x) < F(\sigma, x)$, we know that there are sufficiently small $c > 0$ such that \mathcal{L} will be strongly invasive with respect to x . \square

Taken together, these two theorems show that if adaptive strategies are to invade a population in this setting, the population must be unstable in a strong way.⁸ These results, however, say nothing about whether or not

⁸For instance, population which is currently in evolutionary transition would provide an ideal setting for the invasion of adaptive strategies.

selection will be able to sustain adaptive strategies in stable populations. And, without specifying the character of the response functions we cannot know $\beta(x)$ for any x that includes adaptive strategies, hence we will not be able to analyze such populations.

Informally, there is good reason to suspect that adaptive strategies will not be sustained by selection in this setting. Since adaptive strategies \mathcal{L} are always paying some positive cost here and are restricted to the strategies in ΔS , an individual who adopts the best pure strategy in the support of $r_{\mathcal{L}}$ will presumably do just as well as the adaptive strategies but without paying the cost and hence be able to invade. However, mathematical analysis of this setting cannot be done without substantial restrictions on β and $r_{\mathcal{L}}$. Instead, we will consider a specific example to highlight this point.

2.1 Example: Best Response in Hawk-Dove

Considering a specific example with a very powerful \mathcal{L} reveals the difficulty of overcoming the costs by responding to the population. G will be hawk-dove (shown below). The only adaptive strategy \mathcal{L} will be a strategy that best-responds to the population, or adopts a mixed strategy such that if all learners adopt that strategy the behavior of the population is at the ESS of the game.

Hawk-Dove

	h	d
h	0, 0	3, 1
d	1, 3	2, 2

Using the description of the learning rule \mathcal{L} , we can define β as below.

$$\beta(x) = \begin{cases} (0.5, 0.5) & \text{if } x_h < (x_d + x_{\mathcal{L}}) \text{ and } x_d < (x_h + x_{\mathcal{L}}) \\ ((x_h + x_{\mathcal{L}}), x_d) & \text{if } x_d > (x_h + x_{\mathcal{L}}) \\ ((x_h, (x_d + x_{\mathcal{L}}))) & \text{if } x_h > (x_h + x_{\mathcal{L}}) \end{cases}$$

Where $\beta(x) = (i, j)$ means that the frequency of hawk behavior is i and the frequency of dove behavior is j . According to the corresponding $r_{\mathcal{L}}$, agents adopt h if there is an imbalance of doves, d if there is an imbalance of hawks and a mixed σ that brings the population to the ESS otherwise.

This \mathcal{L} is very powerful, being able to bring mixed populations to the ESS of the game. Never the less, even for very small c ($c = 0.0001$), no computer simulations using the discrete-time replicator dynamic with a random initial population of h , d , and \mathcal{L} resulted in any state other than $x_h = 0.5$, $x_d = 0.5$, and $x_{\mathcal{L}} = 0.0$. Even this powerful learning strategy with a very small cost will not evolve in this setting. Presumably, some additional external environmental change will be required (such as a change in the base game), but in this setting, social interaction alone will not be able to provide the selective pressure for the evolution of adaptive strategies.

2.2 Maynard Smith (1982)

Maynard Smith (1982) in *Evolution and the Theory of Games* suggests that we ought to begin the investigation of the evolution of learning rules by looking at which rules do well when “playing the field” and then apply these rules to repeated contests between individuals (56).⁹ Furthermore, both he and Harley (1981) argue that any evolutionarily stable learning rule (in this population-level setting) will be one that takes the population behavior to the ESS of the game. The model above has given us reasons to doubt both of these claims.

The former is implausible because, as seen above, it seems unlikely that we should see learning evolve due to social interaction at population-level settings. And, if it does evolve, there is reason to doubt that it will be maintained by selection.

The Harley/Maynard-Smith “proof” of the latter claim, that an evolutionary stable learning rule will take the population to the ESS of the game, relied on a learning rule that does take the population to the ESS being able to invade one that does not. Suppose there was a population of \mathcal{L}_j that did not take the population to the ESS of the game, then for some other \mathcal{L}_i that was able to do so, the population of \mathcal{L}_j would seem like a frequency-independent choice setting and \mathcal{L}_i can simply adopt the strategy that is invasive with respect to the behavior of the population.

But, on closer inspection, this reasoning is too quick. There are settings in which a rule for ESSs has no special claim to invasion of populations of learners that are not playing the ESS of the game. Consider an example

⁹For a numerical study of evolutionary stable learning rules in a setting somewhat similar, see Josephson (2008).

game: Rock-Paper-Scissors with an outside option (R-S-P-O, shown below).

R-P-S-O

	r	p	s	o
r	2, 2	4, 0	0, 4	0, 0
p	0, 4	2, 2	4, 0	0, 0
s	4, 0	0, 4	2, 2	0, 0
o	0, 0	0, 0	0, 0	1, 1

In this game there is only one ESS of the game: o . Let population x consist entirely of learners \mathcal{L}_j that take the population to the mixed Nash $(1/3r, 1/3p, 1/3s)$. The average payoff of x is $2 - c_j$ and x cannot be strongly invaded by any \mathcal{L}_i (with $c_i \leq c_j$) since the maximum possible payoff against the behavior of the population is $2 - c_i$ and any strategy used by \mathcal{L}_i also has a maximum possible payoff of $2 - c_i$ against itself. This means that, in the best case, \mathcal{L}_i will be weakly invasive since $F(\mathcal{L}_i, x) = F(x, x)$, and x will be *neutrally stable* with respect to any such \mathcal{L}_i . Hence, \mathcal{L}_i will require a series of mutations to invade and the fact that \mathcal{L}_i is a rule for ESSs does not enable an invasion of x .

At this point, we can draw several conclusions of varying strength. First, learners that respond to populations of individuals will only be able to invade if the population is currently unstable. Second, it seems unlikely that selection in these settings will sustain learners in the long run. Third, these settings do not necessarily suggest anything about the character of the learners that evolve: in particular, they need not be rules for ESSs.¹⁰ As we will see below, when learners are responsive to other individuals there are more possibilities for the invasion and evolution of adaptive strategies.

3 Adapting to Individuals

The extreme opposite of adapting to the population as a whole is adapting to each individual that the agent interacts with. This idealized case may

¹⁰Harley (1981) also attempted to examine the limiting form of ES learning rules and the properties of the relative-payoff-sum learning rule. However, this study has been called into doubt. For a dialogue concerning the validity of these claims see Harley (1983, 1987), Houston (1983); Houston and Sumida (1987) and Tracy and Seaman (1995).

represent a setting where agents have a good memory of specific individuals and have repeated interactions with them. Learners may adopt a different behavior for each individual that they interact with. Since individuals in the model below will have no distinguishing features other than their strategy type, we could also interpret this as “adapting to individual types.”

In this model, we will suppose a large (infinite) population of individuals who are randomly paired to play a repeated game G . In this repeated game, individuals use either a fixed pure strategy, or some *adaptive strategy* \mathcal{L}_i . These adaptive strategies specify a behavior on the first round of the game and provide a method for the individual to modify her behavior as the game proceeds. Each adaptive strategy \mathcal{L}_i will also be assumed to carry an additional fitness cost $c_i > 0$. We will assume that play continues infinitely and that for any two strategies $i, j \in S^L$ there is a well-defined function $\alpha_{ij} : S \times S \rightarrow [0, 1]$ which specifies a probability distribution over pure strategy profiles in G . α_{ij} gives us the long run average frequency of each outcome of game play between an i individual and a j individual. The character of α_{ij} will be left largely unspecified for adaptive strategies and leave the pure strategies unchanged. The constraints on $\alpha_{i,j}$ are as follows:

1. $\alpha_{st}(s, t) = 1$ if $s, t \in S$.
2. $\sum_{s \in S} \alpha_{it}(s, t) = 1$ if $t \in S$ (similarly for i, t reversed).
3. $\sum_{s \in S} \sum_{t \in S} \alpha_{ij}(s, t) = 1$.

To calculate the payoffs of different kinds of learners, we can create a game between adaptive strategies G^L , which will be a type-game derived from G and the character of each adaptive strategy. The payoff for two strategies in S^L will just be the payoff of their long-run interactions in the repeated G .¹¹ The payoff function for G^L can be defined as follows:

$$\pi^L(i, j) = \sum_{s \in S} \sum_{t \in S} \alpha_{ij}(s, t) \pi(s, t).$$

For pure strategies s, t , $\pi^L(s, t) = \pi(s, t)$. But for adaptive strategies, the payoff will be some linear combination of the payoffs. The associated utility and fitness functions can then be defined in the usual way for G^L : $F^L(x, y) =$

¹¹This idealizing assumption allows us to create a fixed payoff matrix for G^L making the analysis much simpler.

$\sum_{j \in S} u^L(x, j)y_j = \sum_{i \in S} \sum_{j \in S} \pi^L(i, j)x_i y_j$. The important thing to note here is that the payoffs to any \mathcal{L} in G^L are just some combination of payoffs in G .¹² We will refer to this model as adaptive strategies “responding to individuals.”

We would like to consider whether or not this setting allows for the possibility of an invasion and evolution of adaptive strategies in a way that the previous model (adapting to the population) did not. The first thing we can show is that there will be a restriction on the sorts of populations that adaptive strategies can invade: adaptive strategies cannot invade stable monomorphic populations.

Theorem 3. *If adaptive strategies respond to individuals, $c > 0$ and x is playing a pure strategy NE of G , then no \mathcal{L} is weakly invasive with respect to x .*

Proof. Suppose that adaptive strategies respond to individuals, $c > 0$ and population x is playing a pure strategy NE of G . Let s be the unique strategy represented in x so $x_s = 1$. It suffices to show that $F^L(\mathcal{L}, s) < F^L(s, s)$. Since \mathcal{L} and s are pure strategies of G^L we have

$$\begin{aligned} F^L(\mathcal{L}, s) < F^L(s, s) & \text{ iff } \pi^L(\mathcal{L}, s) - c < \pi^L(s, s) \\ & \text{ iff } c > \pi^L(\mathcal{L}, s) - \pi^L(s, s) \\ & \text{ iff } c > \sum_{t \in S} \sum_{v \in S} \alpha_{\mathcal{L}s}(t, v) \pi(t, v) - \pi(s, s) \\ & \text{ iff } c > \sum_{t \in S} \alpha_{\mathcal{L}s}(t, s) \pi(t, s) - \pi(s, s) \end{aligned}$$

by definition of π^L and constraint 2 on α . By hypothesis $\pi(t, s) \leq \pi(s, s)$ for all $t \in S$, hence $\sum_{t \in S} \alpha_{\mathcal{L}s}(t, s) \pi(t, s) \leq \pi(s, s)$. Therefore, for any $c > 0$, $F^L(\mathcal{L}, s) < F^L(s, s)$. \square

It follows from Theorem 3 that the extension of any monomorphic ES state x of G will be a monomorphic ES state of G^L . This theorem also means that the strategic settings most conducive to the evolution of adaptive strategies are games without symmetric pure strategy Nash equilibria, because these games cannot have evolutionarily stable monomorphic populations. This class of games includes “competitive” interactions such as

¹²This is a generalization from the modeling framework used by Smead and Zollman (2009) with one exception – here there are no mixed strategies allowed by the framework.

hawk-dove. On the other hand, games with symmetric pure strategy equilibria such as coordination games, the stag hunt, or the prisoner's dilemma will not be conducive to the invasion of learning strategies. With this knowledge in-hand, we turn to showing that there are settings in this model where at least one kind of adaptive strategy can invade.

Theorem 4. *For any game G and any polymorphic population x without adaptive strategies such that for some s represented in x and some $t \in S$, $\pi(s, s) < \pi(t, s)$, if adaptive strategies respond to individuals and c is sufficiently small, then there exists a \mathcal{L} that is strongly invasive with respect to x .*

Proof. Suppose adaptive strategies respond to individuals, x has no adaptive strategies and is polymorphic such that for some s represented in x and some $t \in S$ and $\pi(s, s) < \pi(t, s)$. It suffices to construct a \mathcal{L} such that $F^L(\mathcal{L}, x) > F^L(x, x)$. Let \mathcal{L} play the best response to any pure strategy opponent, so that $\pi^L(\mathcal{L}, v) = \pi(\text{Br}(v), v)$ for each $v \in S$. We now have

$$F^L(\mathcal{L}, x) = \sum_{i \in S} (\pi(\text{Br}(i), i) - c)x_i \quad \text{and} \quad F^L(x, x) = \sum_{i \in S} u(x, i)x_i.$$

$$\begin{aligned} F^L(\mathcal{L}, x) > F^L(x, x) & \text{ iff } \left(\sum_{i \in S} \pi(\text{Br}(i), i)x_i \right) - c > \sum_{i \in S} u(x, i)x_i \\ & \text{ iff } c < \sum_{i \in S} \pi(\text{Br}(i), i)x_i - \sum_{i \in S} u(x, i)x_i. \end{aligned}$$

By definition $\pi(\text{Br}(i), i) \geq u(x, i)$ for all $i \in S$. And, for $i = s$ there is a $t \in S$ such that $\pi(t, s) > \pi(s, s)$, hence $\pi(\text{Br}(i), s) > \pi(s, s)$. By hypothesis, $x_s > 0$, thus $\sum_{i \in S} \pi(\text{Br}(i), i)x_i > \sum_{i \in S} u(x, i)x_i$. Therefore for a sufficiently small $c > 0$, $F^L(\mathcal{L}, x) > F^L(x, x)$ and \mathcal{L} is strongly invasive with respect to x . \square

This theorem suggests that polymorphic populations will provide a good setting for the invasion of adaptive strategies, even if these polymorphic populations are evolutionarily stable with respect to G . Furthermore, the way the invading adaptive strategy is constructed in the proof of Theorem

4 suggests that a traditional “best-response” sort of adaptive strategy may be important for the invasion of learners.¹³ However, we should be cautious with respect to this second point. As Smead and Zollman (2009) show, an adaptive strategy that seeks out a best response cannot be an ESS of game G^L .

3.1 Best Response?

The basic insight from Smead and Zollman (2009) is that, in a broad range of games, when $c > 0$ any \mathcal{L} that best responds to other individuals is susceptible to being outperformed by a strategy that “does the best against its best response.” This is best highlighted by an example. Consider the Hawk Dove game from section 2.1 above, where the three strategies in the population are h , d , and \mathcal{L} . A player using \mathcal{L} will best respond to her opponent and settle on one of the pure-strategy NE against other learners. The type-game G^L can be expressed in matrix form:

Hawk-Dove with Best-Response

	h	d	\mathcal{L}
h	0	3	3
d	1	2	1
\mathcal{L}	$1 - c$	$3 - c$	$2 - c$

As before, we will use computer simulations with the discrete-time replicator dynamic to explore the properties of this game. For $c = 0.01$, every simulation results in a mixture of h and \mathcal{L} . The global dynamics is shown in Figure 1. This picture is very different from the one given by the corresponding game in the adapting to populations model. Here, the adaptive strategy not only invades the population, but is maintained by selection. But, one should also note that this strategy does not dominate the population, and is only sustained in a polymorphic mix.

Given the results of Smead and Zollman (2009) and Theorem 4 above, we may expect that if social interaction alone does provide the selective pressure necessary for the evolution of adaptive strategies, the evolved adaptive

¹³“Best-response” should be read very loosely here. Given the idealizations above, all that is required is that the learners converge on some best response eventually – a property held by most learning dynamics studied in game theory.

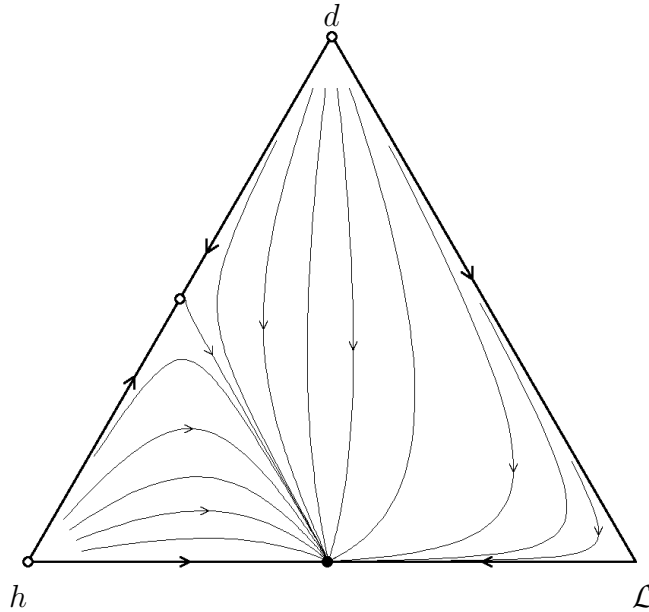


Figure 1: The global dynamics of the Hawk-Dove learning game with Best-Responders and $c = 0.01$.

strategies might not be best responders even though it best responders can cause the initial invasion. As we can see, there are settings where a “non-traditional” adaptive strategy can be evolutionarily stable with respect to the pure strategies in the game. Consider the Prisoner’s Dilemma game with the two pure strategies c and d and an adaptive strategy \mathcal{L}_{cr} that initially chooses a strategy at random (with equal probability) and switches strategies if and only if their opponent outperforms them (call this strategy “Competitive-Response”). The type-game is represented below.

Prisoner’s Dilemma with Competitive-Response

	c	d	\mathcal{L}_{cr}
c	2	0	1
d	3	1	1
\mathcal{L}_{cr}	$2.5 - c$	$1 - c$	$1.25 - c$

In this case, if $c < 0.25$ then \mathcal{L}_{cr} will be an ESS of G^L . Looking at the global dynamics (Figure 2) of this game with the discrete-time replicator

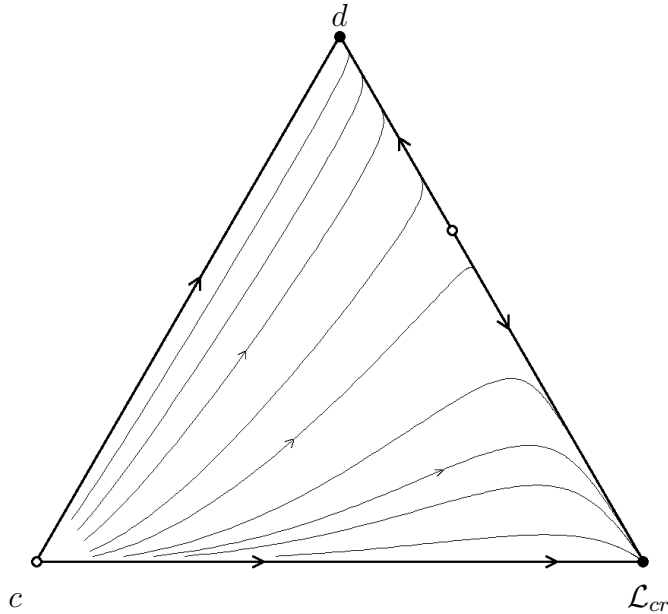


Figure 2: The global dynamics of the Prisoner's Dilemma learning game with Competitive-Responders.

dynamic reveals that even with a large $c = 0.1$, \mathcal{L}_{cr} can have a very large basin of attraction. This suggests that some adaptive strategies may be able to dominate the population even in strategic settings where their initial invasion is difficult. Furthermore, if the same rule is placed in the context of the Hawk-Dove game with Best-Responders, they will form a stable polymorphism of learning strategies. Stable populations that are polymorphic with respect to learners also regularly occur in models of individual and social learning including those of Rogers (1988) and Richerson and Boyd (2000).

Smead and Zollman (2009) may be correct that social interaction is unlikely, by itself, to allow for the evolution of Best-Response type learning rules. However, their analysis does not apply to learning strategies in general. Their results, taken together with the results presented above suggest two tentative conclusions. First, that strategic situations are able to allow for the invasion (although not necessarily complete evolution) of adaptive strategies. Second, if adaptive strategies are to dominate the population they will not be entirely of “traditional” best-response-type learners.

This conclusion may have methodological implications for modeling learning in the social sciences and perhaps the study of strategic plasticity more

generally. For instance, it has been observed by Skyrms (2004) that different learning mechanisms will result in different social contracts. If traditional learning rules should not be expected to evolve in social interaction settings or if we should expect a regular occurrence of stable learning polymorphisms, then selecting a learning rule for modeling purposes may be more difficult (and significant) than previously thought. Furthermore, certain learning rules widely applied in modeling social evolution, such as Herrnstein reinforcement learning,¹⁴ may not be appropriate modeling choices without some additional justification – since these kinds of learning rules may not be expected to evolve in settings of social interaction.

4 Conclusion

We have examined two models of the invasion or non-invasion of adaptive strategies with respect to social interaction: when learners are responding to other individuals and when they are responding to a population. These represent two extremes on a continuum of learning situations. In the middle, learners may be interacting with and/or responding to a subset of the population. Nevertheless, we are able to draw several conclusions from the results of these two models:

1. If certain conditions are met (with respect to the sort of interaction occurring and the composition of the population), social interaction alone is able to provide selective pressure required for the invasion of learning.
2. The ability to adapt to individuals rather than population behavior may be of crucial importance for social interaction alone to provide a setting conducive to the evolution of adaptive strategies.
3. The games that are conducive to the invasion of learners are those associated with an unstable or in some way polymorphic population.

¹⁴This learning mechanism is based on Herrnstein's matching law (Herrnstein, 1970) and is a popular item of study in psychology, evolutionary game theory and philosophy. Beggs (2005) gives an insightful analysis of some of the features of reinforcement learning. More closely related to the role of learning in the evolution of social behavior, Herrnstein reinforcement has very recently been employed in the study of signaling games by many scholars including Argiento et al. (2007); Barrett (2008); Barrett and Zollman (2007); Skyrms (2008); Zollman and Smead (2009).

Upon reflection, this should be somewhat intuitive: these games and populations provide a more varied and complex interaction setting.

4. Although best-response adaptive strategies may be able to invade certain populations, there is evidence from other studies that these same strategies may not be stable. Hence, if learning is to dominate the population in social interaction settings it will likely involve “non-traditional” adaptive strategies.

More generally, the results presented above suggest social interaction may have an important role to play in the evolution of adaptive strategies which is not a straight-forward generalization of complex environments. The interaction of learning rules becomes much more significant: the best way for an individual to learn depends, in a very strong way, on the learning strategies of those they are interacting with. A further and more detailed investigation into the evolution of adaptive strategies will require specifying the adaptive strategies of interest so that the specific behavior (and hence payoffs) can be determined. Furthermore, very little has been said generally about the effects of social interaction on the *character* of the adaptive strategies that might evolve in those settings aside from ability to invade non-adaptive populations. Each of these issues will provide interesting topics for future work.

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. *Theoretical Population Biology* 58, 307–319.
- Argiento, R., R. Pemantle, B. Skyrms, and S. Volkov (2007). Learning to signal: Analysis of a micro-level reinforcement model. *Stochastic Processes and their Applications*.
- Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist* 30(354), 441–451.
- Barrett, J. and K. J. S. Zollman (2007). The role of forgetting in the evolution and learning of language. *Working Paper*.
- Barrett, J. A. (2008). Dynamic partitioning and the conventionality of kinds. *Philosophy of Science* 74, 527–546.

- Beggs, A. W. (2005). On the convergence of reinforcement learning. *Journal of Economic Theory* 122, 1–36.
- Bergstrom, C. T. and P. Godfrey-Smith (1998). On the evolution of behavioral heterogeneity in individuals and populations. *Biology and Philosophy* 13, 205–231.
- Bomze, I. M. and J. W. Weibull (1995). Does neutral stability imply lyapunov stability? *Games and Economic Behavior* 11, 173–192.
- Borenstein, E., M. W. Feldman, and K. Aoki (2008). Evolution of learning in fluctuating environments: When selection favors both social and exploratory individual learning. *Evolution* 62(3), 568–602.
- Borenstein, E. and E. Ruppin (2006). The evolution of imitation and mirror neurons in adaptive agents. *Cognitive Systems Research* 6, 229–242.
- Byrne, R. W. (1997). Machiavellian intelligence. *Evolutionary Anthropology*.
- Byrne, R. W. and L. A. Bates (2007). Sociality, evolution and cognition. *Current Biology* 17, R714–R723.
- Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge University Press.
- Godfrey-Smith, P. (2002a). Environmental complexity and the evolution of cognition. In R. Sternberg and J. Kaufman (Eds.), *The Evolution of Intelligence*. Mahwah: Lawrence Erlbaum.
- Godfrey-Smith, P. (2002b). Environmental complexity, signal detection and the evolution of cognition. In G. M. B. Marc Bekoff, Colin Allen (Ed.), *The Cognitive Animal*, Chapter 18, pp. 135–141. MIT Press.
- Harley, C. B. (1981). Learning the evolutionarily stable strategy. *Journal of Theoretical Biology* 89, 611–633.
- Harley, C. B. (1983). When do animals learn the evolutionarily stable strategy. *Journal of Theoretical Biology* 105, 179–181.
- Harley, C. B. (1987). Learning rules, optimal behaviour, and evolutionary stability. *Journal of Theoretical Biology* 127, 377–379.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior* 15(245-266).
- Hofbauer, J. and K. Sigmund (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.
- Houston, A. I. (1983). Comments on ‘learning the evolutionarily stable strategy’. *Journal of Theoretical Biology* 105, 175–178.

- Houston, A. I. and B. H. Sumida (1987). Learning rules, matching, and frequency dependence. *Journal of Theoretical Biology* 126, 289–308.
- Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson and R. A. Hinde (Eds.), *Growing Points in Ethology*, Chapter 9, pp. 303–317. Cambridge University Press.
- Josephson, J. (2008). A numerical analysis of the evolutionary stability of learning rules. *Journal of Economic Dynamics and Control* 32, 1569–1599.
- Kirchkamp, O. (1996). *Evolution and Learning in Spatial Models*. Ph. D. thesis, Rheinischen Friedrich-Wilhelms-Universität Bonn.
- Kirchkamp, O. (1999). Simultaneous evolution of learning rules and strategies. *Journal of Economic Behavior & Organization* 40, 295–312.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.
- Maynard Smith, J. and G. R. Price (1973). The logic of animal conflict. *Nature* 246, 15–18.
- Richerson, P. J. and R. Boyd (2000). Climate, culture, and the evolution of cognition. In C. Heyes and L. Huber (Eds.), *Evolution of Cognition*, pp. 329–346. MIT Press.
- Rodríguez-Sickert, C., R. Rowthorn, and R. A. Guzmán (2008). The social benefit of slow learners. *unpublished manuscript*.
- Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist* 90, 819–831.
- Sandholm, W. H. (2008). *Population Games and Evolutionary Dynamics*. MIT Press.
- Schlag, K. H. (1998). Why imitate and if so, how? *Journal of Economic Theory* 78, 130–156.
- Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press.
- Skyrms, B. (2008). *Signals: Evolution, Learning, and the Flow of Information*. (book manuscript).
- Smead, R. and K. J. S. Zollman (2009). The stability of strategic plasticity. Working Paper.
- Sterelny, K. (2003). *Thought in a Hostile World*. Blackwell Publishing.
- Taylor, P. and L. Jonker (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* 40, 145–56.

- Thomas, B. (1984). Evolutionary stability: States and strategies. *Theoretical Population Biology* 26, 49–67.
- Tracy, N. D. and J. W. Seaman (1995). Properties of evolutionarily stable learning rules. *Journal of Theoretical Biology* 177, 193–198.
- Weibull, J. W. (1997). *Evolutionary Game Theory*. MIT Press.
- Zollman, K. J. S. and R. Smead (2009). Plasticity and language: An example of the baldwin effect? *Philosophical Studies* (forthcoming).