On Salience and Signaling in Sender-Receiver Games
Partial Pooling, Learning, and Focal Points

Travis LaCroix

Abstract I introduce an extension of the Lewis-Skyrms signaling game, analysed from a dynamical perspective via simple reinforcement learning. In David Lewis' (1969) conception of a signaling game, salience is offered as an explanation for how individuals may come to agree upon a linguistic convention. Brian Skyrms (2010) offers a dynamic explanation of how signaling conventions might arise presupposing no salience whatsoever. The extension of the atomic signaling game examined here—which I will refer to as a *salience game*—introduces a variable parameter into the atomic signaling game which allows for degrees of salience, thus filling in the continuum between Skyrms' and Lewis' models. The model does not presuppose any salience at the outset, but illustrates a process by which accidentally evolved salience is amplified, to the benefit of the players. It is shown that increasing degrees of salience allow populations to avoid sub-optimal pooling equilibria and to coordinate upon conventions more quickly.

Keywords Signaling Games · Salience and Focal Points · Communication Conventions

1 Introduction

To communicate meaningfully, members of a population must cooperate. That is, they must agree upon some convention: one individual ought to use a word or signal to mean one thing if most other individuals in the population use that word or signal in the same way. As such, conventional language use in a population can be modeled with game-theoretic tools. This idea gives rise to

T. LaCroix
University of California, Irvine
Department of Logic and Philosophy of Science
E-mail: tlacroix@uci.edu

the *signaling game*, due to Lewis (1969). To explain how communication conventions might be learned or might evolve, contemporary scholars—following the initial work of Skyrms (1996, 2004, 2010)—have examined families of dynamic evolutionary and learning models which are supposed to capture readily observable communicative phenomena in both human and non-human populations.

Lewis' model presupposes some sort of (shared) salience—e.g., prior agreement, precedent, natural salience, etc.—in order for individuals to come to have 'mutually concordant expectations' about what others might do in a particular situation, thus allowing them to reach a convention by coordinating. In contrast, Skyrms' model presupposes no shared salience. Skyrms shows that, even without shared salience, effective communication can evolve under a variety of starting conditions for a number of different dynamics.

Lewis' model highlights both the social nature of communication and the arbitrariness of the meanings associated with individual signals. From this point of view, linguistic conventions can be understood as stable outcomes of repeated interactions. However, this view turns out to be theoretically untenable for explanatory purposes due to the strength of its assumptions. That is, agents in his model are presupposed to be computationally unbounded insofar as his model makes strong cognitive assumptions about common knowledge and rationality which cannot be appealed to in order to quell skeptical worries about how communication might arise prior to having a language in place with which to agree upon the necessary conventions.

Though Skyrms' model sidesteps these overly strong rationality assumptions, we might note that no shared salience whatsoever is as much of an idealization as salience's being used as the sole explanatory mechanism. Further, little has been said about the possibility for allowing for degrees of salience, or how an evolving signal becomes salient for an individual over time.

In this paper, I introduce an extension of the signaling game, analysed from a dynamical perspective via simple reinforcement learning, which I will refer to as a *salience game*. Specifically, this model introduces a variable parameter into the atomic signaling game, which in turn allows for degrees of salience in the signaling game. As such, this model treats Lewis' and Skyrms' respective models as limiting cases for salience and fills in a continuum of 'game space' between them.

Importantly, the model does not presuppose any shared salience at the outset, but illustrates a process by which accidentally evolved salience is amplified, to the benefit of the players. This sort of approach to modeling allows us to make more realistic assumptions, thus increasing the explanatory value of the results obtained. This is despite the fact that the model itself presupposes very little in excess of Skyrms' sparse model. A practical consequence is that increasing degrees of salience allows individuals to avoid sub-optimal pooling equilibria and to coordinate upon conventions significantly more quickly, as we will see.

Section 2 gives a brief overview of the mechanics of the signaling game and compares the respective models of Skyrms and Lewis. In particular, we will

focus on simple reinforcement learning. Section 3 introduces the notion of a *Schelling focal point* in order to motivate the assumptions that will be made concerning salience. Section 4 introduces the intuitive and formal background for the simulations that were run, with the particular methodology being outlined in Section 4.1. The results of these simulations are given in Section 4.2. With regard to the data obtained from these agent-based simulations, the results are presented with an eye toward average expected payoff (Section 4.2.1), speed of convergence (Section 4.2.2), and pooling equilibria (Section 4.2.3). We find that the higher the probability of salience selection, the faster signaling equilibria are achieved, and that this effect is non-linear in integer frequencies. Finally, Section 5 concludes with a philosophical discussion of this result and the significance of this model.

There are at least two natural interpretations of what this model is supposed to show. On the one hand, we might regard the model as giving a parameter that measures salience out in the world, with the result that an increase in salience (or an increase in awareness of salience for a constant degree of salience) allows for individuals to agree upon conventions more quickly. On the other hand, we might interpret the model as giving a formal analogue for something akin to *demonstration*, at least in a specific kind of communication context. The latter result is closely related to previous work by Barrett and Skyrms (2017) on a constituent of self-assembling games, which they refer to as *cue-reading games*.

2 The Lewis-Skyrms Signaling Game

This section offers a definition and brief historical overview of the signaling game. For a more detailed introduction to simple signaling games and their extensions, see Skyrms (2010).

2.1 Classical Game Theory and the Lewis Signaling Game

The signaling game was introduced by Lewis (1969) to explain how linguistic conventions might arise without there already being a language in place.[1] Using the tools of classical game theory, he formalized interactions in which two players use arbitrary signals to transmit information. This provides a naturalistic account of the emergence of meaning. The simplest case for a signaling game is one in which there are two players (called the *Sender* and *Receiver*), two states of the world (called s_1 and s_2), two possible signals or

---

[1] In this respect, Lewis is responding to skeptical arguments put forth by, e.g., Russell (1922) according to which a new word can be adopted into a language by convention, but a language cannot arise out of a convention. Quine (1967) promotes a similar skepticism that "it is not clear wherein an adoption of the conventions, antecedent to their formulation, consists" (123). See also, Alston (1964).

messages (called m_1 and m_2), and two possible actions (called a_1 and a_2). This is referred to as a 2×2 signaling game.[2]

When agents play a signaling game, nature picks a state of the world at random. In the simplest case, we will assume all states are equiprobable, and so nature is unbiased. The sender observes the state of the world and sends a signal to the receiver. The receiver cannot see which state of the world obtains, but she sees which signal is sent. (As such, this is a game of imperfect information.) The receiver then performs one of the possible actions. We stipulate that a_1 is 'appropriate' in s_1 and that a_2 is 'appropriate' in s_2. Both the sender and receiver receive a payoff of 1 just in case a_i is performed in s_i, and they both receive a payoff of 0 otherwise. The extensive form of the 2×2 signaling game is given in Figure 1.
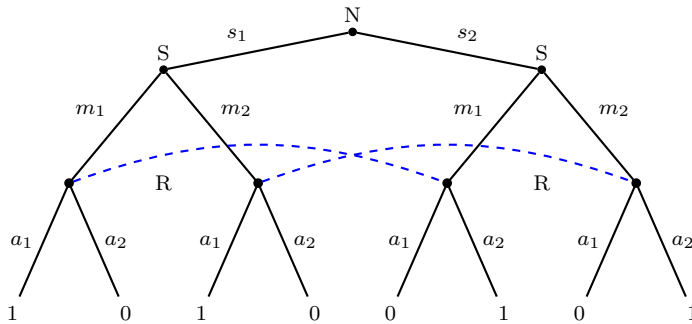

Fig. 1: The extensive form of the simple 2×2 signaling game. Each node denotes a choice point for a given player, and each branch denotes the possibilities available to her at that point. The dotted lines indicate the receiver's information set.

The cases wherein the sender and receiver perfectly coordinate are referred to as the *signaling systems* of the signaling game. The 2×2 game has two possible signaling systems. These are shown in Figure 2. In general, the $n \times n$ signaling game has $n!$ possible signaling systems.[3]

For the symmetric signaling game, where the number of states, signals, and acts coincide, the signaling systems correspond to the strict Nash equilibria of the game. As such, they are supposed to be self-enforcing. Lewis' model thus highlights both the social nature of communication and the arbitrariness of the meanings associated with the individual signals. From this point of view, linguistic conventions can be understood as the stable outcomes of repeated interactions.

---

[2] I will also refer to this as an *atomic* 2-*game*. This terminology is due to Steinert-Threlkeld (2016).

[3] Lewis (1969) proves that, in a signaling game with m states and n signals ($n \geq m$), there are $n!/(n-m)!$ possible signaling systems. For $m = n$, as we have here, it follows that there are $n!$ possible signaling systems.
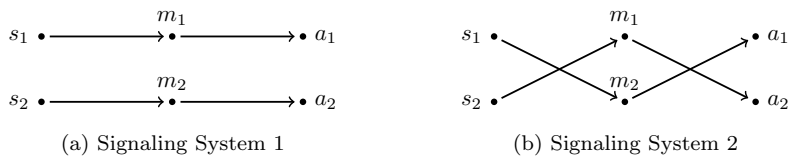
(a) Signaling System 1   (b) Signaling System 2

Fig. 2: The Two Signaling Systems of the 2×2 Game

While it is not explicit in his discussion of language, Lewis (1969) considers repetition implicitly in the following sense. A signaling game is, at bottom, a coordination problem. From Lewis' analysis, a coordination problem can be solved via mutually concordant expectations—high-order expectations about what others might do, what others might expect each other to do, etc. When the same type of coordination problem is repeated, the repetition itself might make some equilibrium salient over and above the others—namely, those that worked in the past. This notion of precedent gives rise to a "metastable self-perpetuating system of preferences, expectations, and actions" (42). We will discuss this further in Section 3.

Even so, Lewis' model of conventions makes strong cognitive assumptions about common knowledge and rationality. In particular, if we assume that agents cannot coordinate via prior agreement, then we will need to build something like *natural salience* into our model. Here, a signal is salient insofar as it is a uniquely and preeminently conspicuous coordination equilibrium (1969, p. 38). Note that this further requires the assumption that the agents in the signaling game notice that one equilibrium is salient, expect the other to notice, etc. Note further that it is unclear exactly how one signaling system should be said to be more salient than any other, without baking the solution directly into the problem. (We will discuss this further in Section 3.) As such, Lewis does not have the resources to explain exactly how a signaling system could come to be adopted as the convention of a population.

2.2 Evolution and Reinforcement Learning

Following the initial work of Skyrms (1996, 2004, 2010), contemporary scholars have examined the signaling game in evolutionary and learning contexts, thus weakening the rationality assumptions inherent in the classical structure upon which Lewis' model is built.

In particular, we will be concerned with a simple reinforcement learning model due to Roth and Erev (1995).[4] Roth-Erev learning is based on Richard Herrnstein's (1970) matching law, under which the probability of an actor's choosing a particular action is proportional to its accumulated rewards. This

---

[4] For more on evolutionary dynamics, see, for example, Hofbauer and Huttegger (2008), Huttegger, et al. (2010), and Skyrms (2010).

in turn is a formalization of the law of effect, due to Thorndike (1905, 1911, 1927).[5]

Simple reinforcement learning in this sense can be understood as so-called 'urn-learning'. We might now imagine that the sender has two urns at her disposal—one of these is labeled 's_1' and the other 's_2'. Each urn contains a ball labeled 'm_1' and a ball labeled 'm_2'. Similarly, the receiver has two urns— which are labeled 'm_1' and 'm_2'—each containing balls labeled 'a_1' and 'a_2'. See Figure 3.

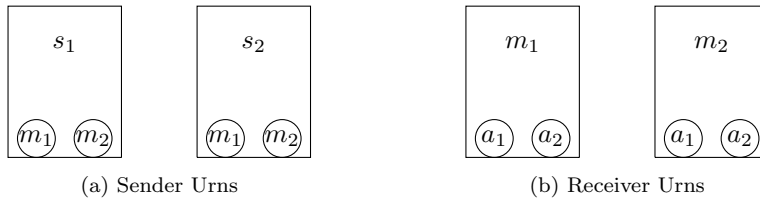(a) Sender Urns   (b) Receiver Urns

Fig. 3: Simple Reinforcement Learning Model

The game now proceeds as follows. Nature picks a state of the world at random. The sender sees the state of the world and selects a ball at random from the corresponding urn—s_1 or s_2. The sender then sends the chosen signal to the receiver, who then selects a ball at random from her corresponding urn— m_1 or m_2. If the act was appropriate for the state, then both the sender and receiver reinforce their behaviour by adding a ball of the same type as the one chosen to the urn from which it was chosen. If the act was not appropriate for the state, then they do nothing.[6]

Instead of a one-off interaction, the game is now repeated. In this way, the addition of a given ball in a given urn makes it more likely that a message [act] will be chosen when a particular state [message] obtains. As such, over time the messages come to carry meaning—or at least information.[7] In spite of its simplicity and relative lack of (e.g., rationality) assumptions, this reinforcement learning model is extremely effective for learning how to signal.[8]

Argiento et al. (2009) show that in the simplest 2×2 signaling game, with equiprobable states, the sender and receiver converge toward one or the other signaling system with probability 1 under this sort of learning dynamic. Therefore, in the limit the players coordinate perfectly. Further, after only 300 plays,

---

[5] For a detailed introduction to reinforcement learning from a computational perspective, see Sutton and Barto (1998).

[6] This is in the simplest case. There are ways of extending the model to include negative reinforcement wherein the sender and receiver discard a ball if it did not lead to a positive payoff. See Skyrms (2010) for an overview of this and several other possible extensions.

[7] Skyrms (2010) uses Kullback-Leibler divergence to quantify the amount of information a given message carries.

[8] On the real-world effectiveness of simple learning, Schultz et al. (1997) show that dopamine neurons in certain areas of primate brains seem to implement a reasonably similar learning procedure. See also Schultz (2004) and Glimcher (2011).

the communicative success rate of the sender and receiver is approximately 0.9, on average (Skyrms, 2010).

Note that we have introduced the dynamic signaling game with an agent-based model and a stochastic dynamics. If we assume that populations are infinite (and so continuous), then it is possible to give, explicitly, a system of differential equations governing the evolution of the population with respect to proportions of strategies—e.g., the replicator dynamic. This gives rise to a deterministic model. The mean-field dynamics of this particular stochastic model (Roth-Erev reinforcement learning) is a version of the (deterministic) replicator dynamic.[9] However, we focus exclusively on the stochastic agent-based learning dynamic.

2.3 Pooling and Partial Pooling Equilibria

In the extended $n \times n$ signaling game it is not certain that a population will arrive at a signaling system. Requiring a communicative success rate of at least 0.8, Barrett (2006) shows that for $n = 3, 4$, failure to converge to a signaling system occurs at rates of 0.096 and 0.219, respectively. For $n = 8$, failure to converge occurs at a rate of 0.594—so, in this case, most runs result in sub-optimal signaling strategies. This is due to the existence of so-called 'pooling equilibria' in the signaling game.[10]

Completely pooling equilibria are strategies in which the sender [receiver] simply ignores the state [message] and performs a particular action. For example, a completely pooling strategy for a sender is to always send m_1 regardless of the state. In this case, signals carry no information. Partial pooling equilibria can occur when the dimension of the signaling game $n > 2$. Figure 4 shows examples of partial pooling outcomes for a 3×3 signaling game.[11]

Pawlowitsch (2008) gives a complete characterization of these partial pooling equilibria and shows (assuming the replicator dynamic) that the partial pooling equilibria in fact have basins of attraction with positive measure. Even though the replicator dynamic in a population setting and simple reinforcement constitute different models, it turns out that learning and evolution are closely related.[12] So, these results are indeed robust.

---

[9] See Beggs (2005) and Hopkins and Posch (2005).

[10] Two caveats ought to be noted here. First, even in the 8×8 game, the players communicate better than chance (where the expected payoff is 0.125). Second, if the communicative success rate is relaxed to 0.75, then failure to converge to a signaling system drops drastically to 0.046. (This highlights the need for a certain attentiveness when building our models.) Further, when states are not equiprobable, pooling equilibria can still be reasonably effective. However, when state probabilities are very asymmetric, the signal may carry little to no information. See Skyrms (2010) and Hofbauer and Huttegger (2008).

[11] See Huttegger et al. (2010).

[12] As was noted above, Beggs (2005) and Hopkins and Posch (2005) show that the mean-field dynamic of Roth-Erev learning is a version of the replicator dynamic. See also Schreiber (2001) for an analysis of the connection between the replicator dynamic and Pólya urns more generally.
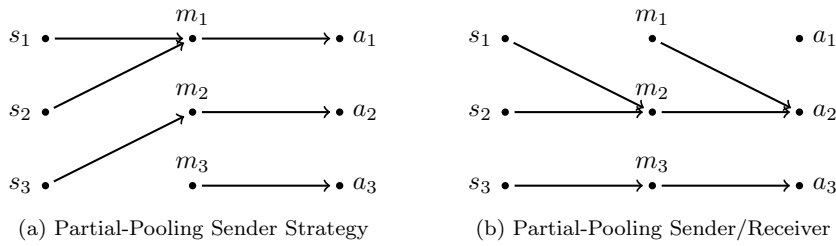
(a) Partial-Pooling Sender Strategy   (b) Partial-Pooling Sender/Receiver

Fig. 4: Two Partial-Pooling Strategies for the 3×3 Game

Skyrms (2010) shows how partial pooling can be avoided by introducing a correlation mechanism to his simple model. Huttegger et al. (2010) show that adding mutation to the (e.g., replicator) dynamic will destabilize total pooling equilibria and also appears to destabilize partial pooling equilibria by reducing the size of the attracting region.[13]

As was noted, salience is also a possible mechanism by which individuals might coordinate (and so avoid pooling equilibria). In this case, we do not change the underlying dynamics of the game, but we change the choice rule for the sender and receiver. The purpose of the subsequent section is to introduce, in more detail, what we mean by salience in order to better understand how it will be utilized in our own model.

3 Salience and Focal Points

Schelling (1960) showed that individual agents can often coordinate their behaviour without communicating by taking advantage of certain salient features of the world. For example, when asked to choose a time and a place to meet in New York City, an absolute majority of respondents chose Grand Central Station for the location, and almost all respondents chose 12:00 noon for the time. Schelling concludes that

> [p]eople can often concert their intentions or expectations with others if each knows the other is trying to do the same. Most situations . . . provide some clue for coordinating behavior, some focal point for each person's expectation of what the other expects him to expect to be expected to do. (57)

Though Schelling's experiments are not rigorous, his general observations have been replicated in a controlled environment.[14]

However, Schelling's focal points require some sort of mutual expectation about what individuals expect others might do in a particular problem. As such, this process requires an individual to engage in high-order reasoning

---

[13] See also Barrett (2006).
[14] See Mehta et al. (1984a) and Mehta et al. (1984b).

about others' expectations about her expectation about others' expectations, etc.

Lewis (1969) writes at length about mutually concordant expectations and this type of high-order reasoning about others' expectations.[15] In particular, for Lewis, coordination requires some sort of salience. Salience may come in several forms, with varying degrees of strength. 'Prior agreement' is perhaps the strongest form of salience, though, as was stated above, this will not help to quell skeptical concerns about the conventionality of language. 'Precedents' constitute another, weaker, type of salience to the extent that two individuals who have come across a certain situation before may simply do what worked last time. Lewis (1969, p. 57) notes that "past conformity to a convention is a basis for common knowledge of a tendency to go on conforming". Finally, natural salience may be a basis for common knowledge, but this will be weaker still insofar as it will generate weak higher-order expectations about what others may do.[16]

One major difference between Lewis' classical signaling game and the evolutionary signaling game is that the latter does not impose strong rationality assumptions. In the Roth-Erev learning model, each signal is initially equally likely for a given state, and each act is initially equally likely for a given signal. As such, there is no salience whatsoever built into this model. Skyrms (2010, p. 8) acknowledges that "there may be signaling systems in nature which got an initial boost from some sort of natural salience", however he wants to assume less in order to strengthen the generality of his model. Thus, his justification for assuming no salience is that this constitutes a "worst-case scenario" wherein signaling systems are *purely* conventional (p. 8).

In considering the worst-case scenario, Skyrms' point is not that salience does not exist in nature, nor that it actually played no role in the evolution of effective communication systems in nature. Rather, the purpose of presupposing no salience whatsoever is to show that even in the worst conceivable case, coordination is still at least possible. In the same way that correlation or mutation may help a population to evolve toward effective communication more quickly, or may help to avoid partial pooling, so too may some degree of salience.

In particular, some level of salience between zero in Skyrms' model, and the over-reliance on salience in Lewis' model, deserves some attention insofar as no salience whatsoever is as much of an idealization as salience being used as the sole explanatory mechanism. Rather than making any presuppositions about

---

[15] As a paradigm case, Lewis writes that

I may reasonably expect you to realize that, unless I already know what you expect me to do, I may have to try to replicate your attempt to replicate my reasoning. So I may expect you to try to replicate my attempt to replicate your attempt to replicate my reasoning. So my own reasoning may have to include an attempt to replicate your attempt to replicate my attempt to replicate your attempt to replicate my reasoning. And so on. (27-8)

[16] Schelling (1960) notes that determining focal points of a coordination problem often relies more on imagination than logic.

salience in this model, we ask the questions: *How much salience is sufficient to avoid partial pooling or to allow communication to evolve more quickly? How much more efficient does salience make the evolution of communication? How might salience itself evolve*—i.e., *how might an evolving signal* become *salient?*

Little has been said about the possibility for some salience, and it is unclear whether the classical game-theoretic approach even allows for such a concept. However, it is an advantage of the evolutionary approach to signaling games that we can make sense of this notion and subsequently investigate it. As a result, we are able to examine how agents might be expected to exhibit behaviour which takes advantage of this salience by chance. We explore such an extension in the subsequent section.

4 The Model

The motivation underlying this model is to supplement the atomic signaling game with a salience parameter to fill in the continuum between Lewis' and Skyrms' respective idealizations.

The model described above for the atomic signaling game under simple reinforcement learning allows the sender and receiver to take advantage of evolved salience as the game progresses and actions are reinforced. For example, suppose a signaling game has progressed in such a way that the sender's probabilities are as shown in Figure 5. In this case, the probability that the

| | | Messages | |
		m_1	m_2
States	s_1	0.500	0.500
	s_2	0.500	0.500

(a) Initial Propensities

| | | Messages | |
		m_1	m_2
States	s_1	0.999	0.001
	s_2	0.003	0.997

(b) Evolved Propensities

Fig. 5: Sender Probabilities

sender will send m_1 given that nature chooses s_1 is 0.999, and the probability that she will choose m_2 given s_2 is 0.997. Hence, we might say that m_1 has become salient for the sender when s_1 obtains. Similarly, m_2 has become salient for the sender when s_2 obtains. In this sense, the sender takes advantage of a salience that was not initially present, but has evolved through the learning process. That is to say, a precedent has been set through simple trial-and-error.

The *salience game* works exactly like the atomic signaling game, except it includes a salience parameter which amplifies the accidentally evolved precedents that are already present in the atomic game, to the advantage of the sender and the receiver.

As with the atomic 2×2 signaling game, in the 2×2 salience game there are two states of the world (called 's_1' and 's_2'), two possible signals or messages (called 'm_1' and 'm_2'), and two possible actions (called 'a_1' and 'a_2').

We further stipulate a *salience parameter*, μ, which can be interpreted as a probability or a frequency. On a normal round of game-play (with probability $(1 - \mu)$), the sender and receiver play as was described in Section 2.2: Nature picks a state of the world at random, the sender chooses an initially random action, and both the sender and receiver reinforce their choices just in case they coordinate. As the game continues, the probabilities with which they choose their respective messages or actions change proportional to the accumulated rewards for coordination.

However, with some probability (the salience parameter, μ), the sender [receiver] uses a different choice procedure. In this case, the sender [receiver] chooses randomly from the set of most-likely signals [actions] given the particular state [signal]—i.e., the sender [receiver] chooses from a set of signals [actions] that are most salient in the given context. As such, with probability $(1 - \mu)$, the choice rule is given by the normal Roth-Erev reinforcement learning procedure—choosing stochastically based on their respective probabilities, which are proportional to the accumulated rewards. However, with probability μ, the sender and receiver use a "greedy" choice rule to pick whatever action is most probable (i.e., most salient). If no unique action is most salient, they ignore less-probable actions and randomize over the remaining actions.

To clarify what is happening in the salience case, let $M_{s_j} = \{m_1, \ldots, m_n\}$ be the set of possible signals for a given state, s_j, from which the sender chooses. (Note: symmetric remarks will apply for the set $A_{m_j} = \{a_1, \ldots, a_n\}$ from which the receiver chooses an action for a given message, m_j.) For each given state, there is an associated vector of probabilities,

$$P_{(m_i|s_j)} = \langle p(m_1 \mid s_j), \ldots, p(m_n \mid s_j) \rangle,$$

that a given signal is chosen in that state, with $\sum_{i=1}^{n} p(m_i|s_j) = 1$. The key distinction between the atomic choice function and the salience choice function is as follows: in the atomic case, the sender chooses from the entire set, M_{s_j}, with the likelihood of choosing a particular signal, m_i, being given by the associated weight, $p(m_i \mid s_j)$, for that signal in state s_j—again, this probability is proportional to the accumulated reward thus far. However, in the salience case, the sender chooses rather from the following set:

$$M'_{s_j} = \arg\max_{m_i} P_{(m_i|s_j)}.$$

For the n-game with $n > 2$, there are three distinct cases to consider, depending on the cardinality of the set M'_{s_j}: (i) $|M'_{s_j}| = 1$, (ii) $1 < |M'_{s_j}| < n$, and (iii) $|M'_{s_j}| = n$.[17] In the singleton case, (i), there is only one signal to choose, and so the sender chooses that signal. This might be understood as a particularly salient signal given the state, though we will say more about interpretation in Section 5. The other extreme is case (iii): since no signal is most likely (all signals are equiprobable, and $M' = M$), there is no salience to

---

[17] Note, if $n = 2$, then case (ii) is excluded.

rely upon, and therefore the sender must choose randomly between all possible signals. In this case the sender's choice procedure is *de facto* atomic.[18]

In case (ii), the sender does not have a uniquely salient signal, but ignores unlikely signals. This, in effect, reduces the dimension of the game for a given round. Since the probability of arriving at a signaling system is greater for lower-dimensional games, this small degree of salience can have a significant impact on the eventual outcome of the game.

Table 1: Three Possibilities for the Sender in the Salience Game

	m_1	m_2	m_3	
s_1	0.33	0.33	0.33	Case (iii)
s_2	0.20	0.40	0.40	Case (ii)
s_3	0.80	0.05	0.15	Case (i)

By way of a concrete example, consider the possible sender strategy for a 3×3 game by some given play of the game, shown in Table 1. In this example, we have

$$M'_{s_1} = \{m_1, m_2, m_3\},$$
$$M'_{s_2} = \{m_2, m_3\}, \text{ and}$$
$$M'_{s_3} = \{m_1\},$$

corresponding to cases (iii), (ii), and (i), respectively. With probability $(1-\mu)$, she chooses (atomically) according to the accumulated reward. Thus in s_1, she chooses uniformly between m_1, m_2, and m_3. In s_2, she chooses m_1 with probability 0.2, she chooses m_2 with probability 0.4, etc. Finally, in s_3, she chooses m_1 with probability 0.8, she chooses m_2 with probability 0.05, etc.

With probability μ—i.e., the salience case—she greedily chooses the most probable or most salient message. In s_3, m_1 is particularly salient, and so the sender chooses this signal. In s_2, m_1 is unlikely. However, there is no uniquely salient signal. Thus, the sender chooses from the set $\{m_2, m_3\}$, with each message being equiprobable. As such, it is as if the sender were playing a 2×2 game on this particular round. Finally, in s_1 no salience has evolved yet, and so the sender effectively uses the atomic choice procedure.

In case (iii), the sender is a *de facto* atomic sender. This highlights the fact that no salience here is presupposed at the outset. Rather, the salience of a particular signal [act] given a particular state [signal] changes as the game-play proceeds. We do not stipulate that, e.g., signal i is salient in state j: the salience evolves as a function of the actor's evolving individual choices and so *becomes* salient as time goes on. The model illustrates a particular process through which accidentally evolved salience is amplified. In the atomic signaling game, we may also understand the process as an evolution of some

---

[18] Technically, she chooses according to her accumulated reward; however, in this case all messages are equiprobable, so this is tantamount to randomizing over all available messages.

salience. In particular, early random salience is amplified by way of increasing weights in general. However, the process of amplification here is stronger.

Before outlining the results for the salience game, we note that the reinforcement learning procedure here is effectively the same as Roth-Erev reinforcement learning, as was outlined in Section 2.2, to the extent that there are no additional presuppositions going in to the model. The only difference between the atomic game and the salience game is that with probability μ, the choice rule is different.

The choice rule that we have described is taken as an analogue for salience to the extent that "most probable" is understood as "most salient". There are, of course, other parametric choice functions that might be used here. For example, Blume et al. (2002) introduce a softmax function as an alternative choice rule for Roth-Erev reinforcement learning. This approach defines an *exponential response rule*, where the probabilities are proportional to the exponential of accumulated rewards rather than the accumulated rewards themselves. Generally, the accumulated rewards are multiplied by a constant, λ—the reciprocal of λ is sometimes called the "temperature". When λ is 0, the reciprocal is infinite, so the choice rule is uniform among the possibilities. Similarly, when λ is arbitrarily large, the reciprocal approaches 0, and the act with the largest accumulated reward is chosen with the highest probability.

Therefore, as $\lambda \to \infty$, the softmax choice rule thus defined mirrors our greedy "salience" choice rule. In this sense, the choice rules are similar under a particular context; however, though softmax performs similarly in certain contexts to the choice rule we have defined here, it is less clear how we might interpret this choice rule in the context with which we are concerned—namely, modeling salience.[19] In the next section, we present our simulation results to show the effect of salience.

4.1 Method

Numerical simulations of simple reinforcement learning with only positive reinforcement payoffs equal to 1 were run. Since we are particularly interested in the short-term evolution of the game, each run consists in $10,000$ individual plays of the game, and we take an average of the results of 1000 runs. Though this is a reasonably low number for a simulation of this sort, it was chosen to make apparent the difference in speed of learning between the atomic game and varying amounts of salience. In order to get a better picture of the limiting

---

[19] That is to say, as a bit of mathematical machinery, the softmax choice rule might have arrived at the same results; however, the extent to which these results are explanatory (at least in a 'how possibly' sense) are going to be limited by our ability to interpret this choice rule as having some natural analogue. See Blume et al. (2002) and Skyrms' (2010) discussion for further explication of softmax choice.

behaviour of these simulations, a significantly larger number of trials per run would be necessary.[20]

In order to get a reasonably complete picture of how salience affects the expected payoffs and convergence toward signaling, we examine cases in which the salience parameter is $\mu = 0.00, 0.01, 0.025, 0.05, 0.10, 0.15, 0.20, 0.25, 0.50$ and 1.00. Note that when $\mu = 0.00$, the model is identical to Skyrms' model, where there is zero initial salience. When $\mu = 1.00$, the salience is not assumed, but an initial success sets a strong precedent which individuals then rely upon for all future plays. As such, this is a version of Lewis' model, which relies upon precedents rather than natural salience. In order to examine the robustness of the results across more complex games, we look at several dimensions of the signaling game, with $n = 2, 3, 4, 5, 6, 7$ and 8.

4.2 Results

The following subsections display graphical and tabular data of the results of the simulations for a variety of statistics which may be of interest. In particular, we examine the average expected payoff for 1000 runs by the end of the $10,000$ plays for each parameter in order to get a general measure of performance. We also compare the speed of convergence to a cutoff parameter between the atomic game and the salience game. Finally, we compare the proportion of runs that fail to reach this cutoff parameter between the atomic and the salience game by examining the pooling effects for varying degrees of salience.

4.2.1 Average Expected Payoff

The average expected payoffs for 1000 runs across the parameters examined is summarized in Table 2. A graphical display of these data is shown in Figure 6.

As can be seen, an increase in salience results, in general, in an increase in expected payoff for the players. However, this is not necessarily true in lower-dimensional cases. In particular, when $n = 2$, we see that the salience parameter for $\mu = 0.05, 0.10, 0.15, 0.20, 0.25$ has lower expected payoff than the atomic game ($\mu = 0.00$). However, the difference here is negligible: it is, on average, smaller than 0.0030. As such, the slight clustering above and below the atomic expected payoff is likely a result of the low number of trials that were run. The expected payoff is sufficiently close in this case that we can reasonably say, in general, that allowing for salience is at least as efficient as in the 2-game. Similarly, with the exception of the cases where $\mu = 0.01, n = 3$ and $\mu = 0.025, n = 4$, respectively, a monotonic increase in salience results in a monotonic increase in expected payoff for $n = 3, 4, 5, 6, 7, 8$.

In fact, we can prove both of these observations analytically. In order to do so, we introduce the following formal definitions.

---

[20] The simulations were run with Python 2.7, and the data from the simulations was compiled and interpreted with MatLab. The code can be found here: github.com/travislacroix/demonstrator-game.

Dimension of Game, n

	2	**3**	**4**	**5**	**6**	**7**	**8**
0.00	0.9866	0.9413	0.9080	0.8819	0.8624	0.8417	0.8172
0.01	0.9867	0.9359	0.9169	0.8907	0.8679	0.8504	0.8319
0.025	0.9879	0.9421	0.9163	0.8946	0.8750	0.8599	0.8469
0.05	0.9847	0.9453	0.9218	0.9049	0.8840	0.8784	0.8685
0.10	0.9843	0.9507	0.9296	0.9189	0.9158	0.9119	0.9002
0.15	0.9821	0.9547	0.9418	0.9434	0.9351	0.9359	0.9257
0.20	0.9824	0.9703	0.9616	0.9558	0.9531	0.9509	0.9415
0.25	0.9850	0.9728	0.9704	0.9719	0.9656	0.9609	0.9564
0.50	0.9986	0.9967	0.9950	0.9924	0.9891	0.9848	0.9802
1.00	0.9996	0.9988	0.9976	0.9960	0.9940	0.9915	0.9887

Salience Parameter, μ

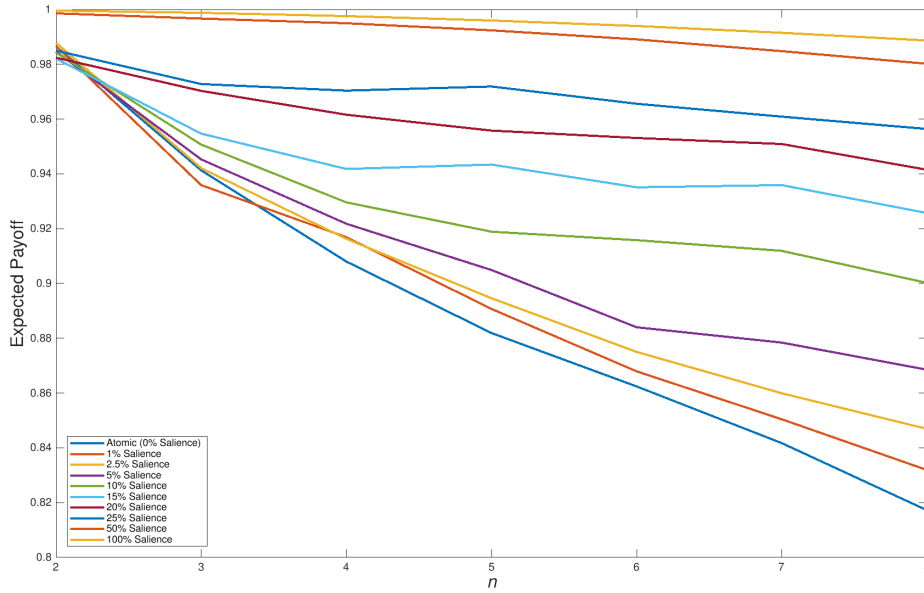Table 2: Average Expected Payoff Across Parameters

Fig. 6: Average Expected Payoff Across Parameters

Definition 1 The communicative success rate (expected payoff) for the atomic n-game is given by

$$\pi(\sigma, \rho) = \sum_{s \in S} P(s) \sum_{a \in A} u(s,a) \cdot \left(\sum_{m \in M} \sigma(s)(m) \cdot \rho(m)(a) \right)$$

where $s \in S, m \in M$, and $a \in A$ are states, messages, and acts (respectively), with $|S| = |M| = |A| = n$; the sender strategy, σ, is a function mapping states to a probability distribution over messages; the receiver strategy, ρ, is a function mapping messages to a probability distribution over acts; $P(s)$ is a probability distribution over states (uniform in our case); and $u(s, a)$ is a utility function—the Kronecker delta, $\delta_{s,a}$, in this case.[21]

Recall that the choice rule for the salience n-game is differentiated from the atomic n-game in the following way: When there is no uniquely salient message given the state for the sender, she chooses a message proportional to that state's accumulated rewards.[22] (Note that similar remarks hold for the receiver choosing an act based on a message.) When there is a subset of messages that are equiprobable, the sender randomizes over the subset, rather than the entire set of messages. Finally, when there is a uniquely salient message (one message that is most probable), the sender simply picks that message.

For simplicity, let us assume the probability distribution over each message has a uniquely salient state (whose probability is higher than all the other messages, given that states).[23] The sender and receiver always coordinate when they choose greedily in this case, since their probability distributions change at the same time. Thus, we can define (under this assumption) the communicative success rate for the salience n-game as follows.

Definition 2 The communicative success rate for the salience n-game is given by

$$\pi^{\mu}(\sigma, \rho) = (1 - \mu)\pi(\sigma, \rho) + \mu$$

That is, the sender and receiver obtain expected payoff $\pi(\sigma, \rho)$ with probability $(1 - \mu)$—i.e., when they play the game as usual—and they coordinate perfectly with probability μ—i.e., when they play the salience game.

With these definitions in hand, we can easily prove the following two propositions.

Proposition 1:
The communicative success rate for the salience n-game is at least as successful as the atomic n-game, with equality just in case $\mu = 0$ or $\pi(\sigma, \rho) = 1$.

---

[21] See Huttegger (2007) or Steinert-Threlkeld (2016) for a detailed formal characterization of the signaling game in general.

[22] The case where all messages are equiprobable amounts to randomizing over all the messages, which would mean that the expected payoff is equivalent to the communicative success rate $\pi(\sigma, \rho)$ for the atomic n-game, as given in definition 1 above.

[23] Note that this is a fairly strong simplifying assumption. However, this allows for a significantly more concise proof. It would be fairly straightforward to accommodate "in between" cases in the definition for the communicative success rate. However, this would have to be done recursively on the dimension of the game for a fully general proof. Thus, for space considerations we help ourselves to this assumption.

proof:

Assume that $\mu \geq 0$. Then it follows that $\mu(1 - \pi(\sigma, \rho)) \geq 0$. This is true because $0 \leq \pi(\sigma, \rho) \leq 1$, given our utility function, $u(s, a)$, so it follows that $0 \leq (1 - \pi(\sigma, \rho)) \leq 1$. If we distribute the μ term, we see that $\mu - \mu(\pi(\sigma, \rho)) \geq 0$. Adding $\pi(\sigma, \rho)$ to both sides of this inequality gives us $\pi(\sigma, \rho) - \mu(\pi(\sigma, \rho)) + \mu \geq \pi(\sigma, \rho)$. Finally, factoring out the $\pi(\sigma, \rho)$ on the left side of the inequality gives us $(1 - \mu)\pi(\sigma, \rho) + \mu \geq \pi(\sigma, \rho)$. However, by Definition 2, we have $(1 - \mu)\pi(\sigma, \rho) + \mu := \pi^{\mu}(\sigma, \rho)$. Therefore, by substitution, we have

$$\pi^{\mu}(\sigma, \rho) \geq \pi(\sigma, \rho)$$

with equality just in case $\mu = 0$ or $\pi(\sigma, \rho) = 1$, which is what we wanted to show. $\qquad\square$

Proposition 2:

The communicative success rate $\pi^{\mu}(\sigma, \rho)$ is a monotonically increasing function of μ.

Proof:

Assume that $\mu' > \mu$. Then it follows that $\mu'(1 - \pi(\sigma, \rho)) > \mu(1 - \pi(\sigma, \rho))$. This is true because $0 \leq \pi(\sigma, \rho) \leq 1$, given our utility function, $u(s, a)$. It follows that $0 \leq (1 - \pi(\sigma, \rho)) \leq 1$. Distributing the μ' term on the LHS of the inequality, and the μ term on the RHS of the inequality gives us $\mu' - \mu'\pi(\sigma, \rho) > \mu - \mu\pi(\sigma, \rho)$. Adding a $\pi(\sigma, \rho)$ term to both sides of the inequality (and reordering terms) gives us $\pi(\sigma, \rho) - \mu'\pi(\sigma, \rho) + \mu' > \pi(\sigma, \rho) - \mu\pi(\sigma, \rho) + \mu$. By factoring out the $\pi(\sigma, \rho)$ on the left- and right-hand sides of the inequality, we obtain $(1 - \mu')\pi(\sigma, \rho) + \mu' > (1 - \mu)\pi(\sigma, \rho) + \mu$. Thus, by definition, we have

$$\pi^{\mu'}(\sigma, \rho) > \pi^{\mu}(\sigma, \rho)$$

for $\mu' > \mu$, which is what we wanted to show.

$\qquad\square$

Even if we do not assume that each state has a uniquely most-probable signal, it should be clear that because of the reduction in dimensionality, the expected payoff is increased; however, the argument would be more complicated in this case.

4.2.2 Speed of Convergence

We also examine the speed at which the various games reach or exceed a particular cutoff parameter—in this case, 0.90—by the end of the run (10,000 trials). The run failure rate for 1000 runs across the parameters examined is summarized in Table 3.

Dimension of Game, n

	2	**3**	**4**	**5**	**6**	**7**	**8**
0.00	0.033	0.173	0.331	0.498	0.624	0.741	0.860
0.01	0.032	0.186	0.296	0.464	0.604	0.723	0.818
0.025	0.027	0.169	0.302	0.440	0.573	0.685	0.771
0.05	0.036	0.164	0.284	0.401	0.514	0.607	0.664
0.10	0.032	0.138	0.254	0.338	0.390	0.429	0.511
0.15	0.036	0.127	0.202	0.224	0.287	0.302	0.365
0.20	0.038	0.082	0.130	0.166	0.194	0.210	0.276
0.25	0.028	0.074	0.094	0.094	0.127	0.155	0.164
0.50	0.001	0.004	0.005	0.008	0.015	0.025	0.033
1.00	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Salience Parameter, μ (row labels)

Table 3: Run Failure Rates Across Parameters (0.9 Cutoff)

The cutoff parameter chosen here is 0.9. For theoretical purposes, this is excessively high, and the relevant cutoff parameter is going to vary depending on the dimension of the game. However, 0.9 was chosen simply for consistency of results across the salience parameters—i.e., for ease of comparison. The numbers in Table 3 do not mean that, e.g., 33 of the atomic 2-game runs failed to converge to a *signaling system*, but simply that by the 10,000th trial they failed to have an expected payoff of at least 0.9. In this way, the data in Table 3 speaks to the speed at which the players learn to communicate efficiently. As we can see, an increase in salience generally translates to an increase in speed of learning how to signal. To illustrate this clearly, we can compare, for example, the failure rates in the 8×8 game. In the atomic game, very few of the runs have achieved effective communication. However, an increase in salience by $0.05, 0.10$ and 0.15 results in approximately $150\%, 250\%$ and 350% increases in successes, respectively. As such, salience allows individuals to learn how to effectively communicate more quickly—again, with an increase in frequency of salience resulting, in general, in an increase of speed.

Unsurprisingly, we find quantitative similarities in the lower dimensional cases as we did with the expected payoffs of these games (4.2.1). That is, with the exception of the cases where $\mu = 0.01, n = 3$ and $\mu = 0.025, n = 4$, respectively, a monotonic increase in salience results in a monotonic decrease in failures for $n = 3, 4, 5, 6, 7, 8$.

In particular, we may note that, while an increase in salience (generally) translates to a decrease in failures across parameters, the difference between salience rates of $\mu = 0.00$ and $\mu = 0.05$ is significantly higher, for example, than an increase in salience rates between $\mu = 0.20$ and $\mu = 0.25$. So, there is not a linear relation between differences in salience rates and differences in success rates. However, in general, we ought to note that higher dimensional games tend to have significantly higher increases in success. For example, for a 5% increase in salience, $\mu = 0.05$ sees a $7.03\%, 19.32\%$, and 29.26% increase in success for $n = 4, 5, 6$, respectively, as compared with $\mu = 0.00$. As such,

in several of these simulations we see that salience is *more* effective in more complex games than it is in less complex games.

We analyse the variance with a factor of $\mu \times n$ using a linear regression model to determine whether there is a statistical interaction effect between the variables for salience and dimension. The results in Figure 7 show the effect of increasing salience on failure rates does appear to depend in some way upon the dimension of the game.
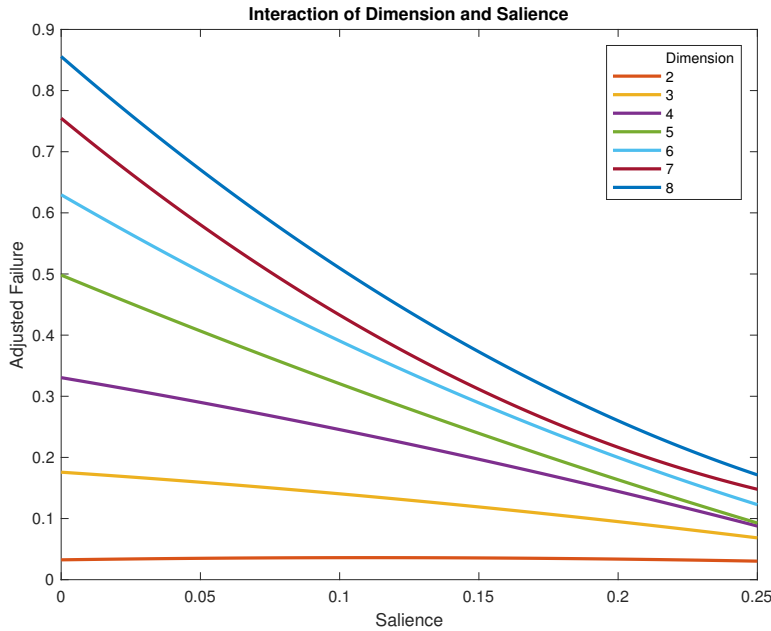

Fig. 7: Interaction Effect Between Salience and Dimension

We might wonder why this is the case. Note that a small increase in dimension corresponds to a significant increase in the strategy space of the game.

For an $n \times n$ signaling game, there are n^{2n} combinations of strategies. Thus, there are 713 more strategy combinations in the 3×3 game than there are in the 2×2 game, but there are $280,796,753,637,807$ more combinations of strategies in the 8×8 game than there are in the 7×7 game. As such, it makes sense that salience would have a larger effect on higher-dimensional games: we said above that the salience choice function effectively reduces the dimension of the game for that particular choice. This corresponds to a significant reduction in the strategy space for the player. Thus, coordination happens more quickly.[24]

---

[24] Note that, under a deterministic dynamics with a continuous population, one can give a measure of the basins of attraction for the partial pooling equilibria. However, we have

4.2.3 Pooling Equilibria

The cutoff parameter that one chooses is important insofar as the expected
payoff of reasonably effective pooling equilibria is going to vary depending on
the dimension of the game. For example, in the atomic 3-game, the optimal
solution is one of the six signaling systems that, in the limit, has an expected
payoff arbitrarily close to 1. However, there will be some pooling equilibria
such that, if the expected payoff by the ith trial exceeds the theoretical payoff
of these pooling equilibria, then the players have already "escaped" the poly-
morphic trap, and will eventually converge to a signaling system. When states
are equiprobable, the most efficient pooling equilibrium in, e.g., the atomic
3-game, has an expected payoff of 0.666667. As such, if the expected payoff by
the ith trial is, e.g., 0.75, then the population will eventually arrive at a signal-
ing system with an expected payoff of 1. Note that there are several different
polymorphic traps that have this expected payoff—however, for mathematical
purposes, they are equivalent.

However, if the dimension of the game is higher, then there are going to
be a larger number of pooling equilibria,[25] and the most effective pooling
equilibria are going to have an expected payoff closer to the optimal solution.
For example, the expected payoff of the most optimal sub-optimal equilibrium
in the atomic 8-game is 0.875. Thus, if a run exceeds this payoff by the ith
trial, then it is convergent toward an optimal solution.

Figure 8 compares the threshold of sub-optimal pooling equilibria for the
actual expected payoff of the 1000 runs of each game with $n = 6$. The vertical
dashed lines indicate the expected payoff for sub-optimal pooling equilibria.

Here, we note that as μ increases, there is more weight on signaling systems
and less weight on partial pooling equilibria. Furthermore, given that the time-
scale is the same for each combination of parameters, we can compare the rate
at which the individuals escape these pooling equilibria. As can clearly be seen,
by the $10,000$th trial significantly more runs end up with a higher expected
payoff, and as the salience increases the speed with which we arrive at the
respective payoffs increases.

The rate of convergence toward equilibria shown here is reasonably quicker
across the variety of salience rates, with an increase in salience resulting (gen-
erally) in an increase in efficiency. Again, since we only examined short-term
results for 10,000 iterations per run, it is not necessarily the case that the
runs which appear to be converging to sub-optimal equilibria will stay there.

focused here on an agent-based reinforcement learning model, which is stochastic, and so
no such measure exists for the way this model was presented. Given the relation between
the mean-field dynamic for Roth-Erev reinforcement learning and the replicator dynamic,
it would be theoretically possible to obtain such a measure for the salience game and so to
show analytically whether and how these basins of attraction decrease; however, this would
be fairly costly with regard to space considerations. We still believe the results shown in
this section serve as illuminating gesture in that direction.

[25] In fact, this number will grow super-exponentially as the dimension of the game in-
creases, in the same way that the number of signaling systems in a particular game grows
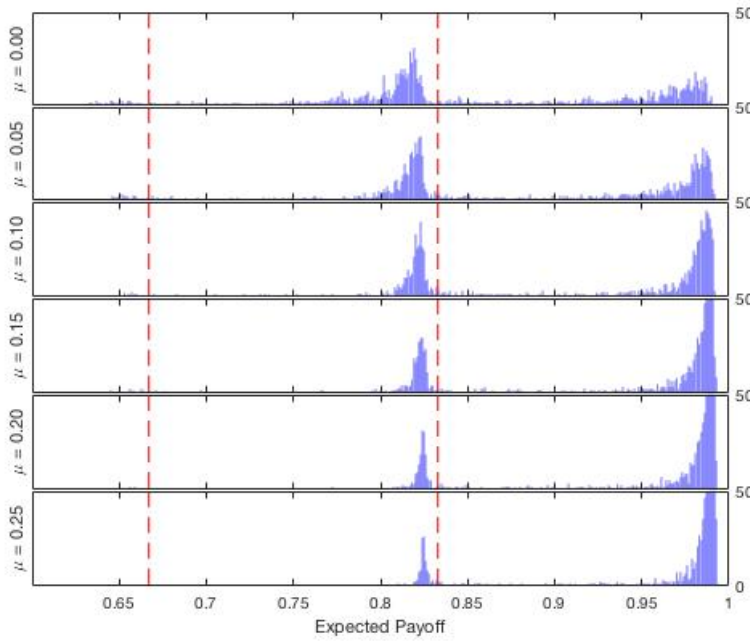super-exponentially as the dimension of the game increases.

Fig. 8: Pooling Around Optimal and Sub-Optimal Strategies for a 6×6 Signaling Game

Similar results hold across the various dimension parameters, with the intuitive outcome that more salience tends the data toward the right of the graph, and similarly for lower dimensional games, where convergence happens with greater certainty.

5 Concluding Remarks

We might note the following interpretation of what we mean by salience in this paper. It appears that the type of salience to which Schelling refers, in his discussion of *focal points*, is an *objective* salience. To the extent that Lewis is couching his discussion of salience in terms of Schelling's work (and similarly, to the extent that Skyrms is discussing salience in terms of Lewis' work), this objective feature of salience seems to be inherited by each of these models.

However, because of the asymmetric information inherent to the signaling game—the sender knows the state of the world, but the receiver does not—it appears that the salience we are concerned with here is a *psychological* (or *subjective*) rather than an *objective* notion of salience. Namely, the features of a particular message that may seem salient to the sender, given her knowledge of the state of the world, might not be salient to the receiver—e.g. a signal

with *context* may seem particularly salient for that context, but without the additional contextual information, it may seem random.

In the terms of our model, internal salience is created for both the sender and receiver by the rewards generated by their past behaviour. Due to the informational asymmetry between the sender and receiver, we might think that the salience available to the sender is not available to the receiver. There is indeed a non-zero probability that the sender and receiver will evolve *different* saliences early on in the game.[26] However, we note that this is a game of *pure* coordination. As such, most of the time, the sender and receiver will evolve a shared salience, since they only reinforce their behaviour when they coordinate. Furthermore, this is a self-correcting process—they cannot evolve "opposite" saliences because they would then fail to coordinate and so not reinforce their behaviour (i.e., this is unstable).[27]

As such, the asymmetry of information in the signaling game makes it appear that we are concerned with a subjective salience as opposed to an objective salience (and so diverge from the usual interpretation of salience in the sense of Schelling). However, since the game is one of pure coordination, the subjective saliences that get evolved are often going to be shared. As a matter of mechanics, I do not believe the distinction is a significant one.

Furthering this point, we might regard the model as giving a parameter that measures salience "out in the world", with the result that an increase in salience allows for individuals to agree upon conventions more quickly. This is a metaphysical, and so *objective* understanding of salience, wherein the matter of fact about whether a particular state of the world has some salience is going to vary.[28] On the other hand, we might interpret the variable salience parameter as modeling a constant degree of salience in the world, about which the agents in the game are variably aware—or, likewise, we might say the agents are only intermittently motivated to attend explicitly to what might be most mutually salient. This is an epistemic interpretation of salience, wherein the matter of fact is that a given state of the world is always salient, but actors do not always notice. This is more akin to the subjective notion of salience. Again, though there may be some interesting philosophical consequences lurking about, I do not believe anything significant turns on this distinction for our purposes here.

---

[26] For a concrete example: suppose on the first round, choosing stochastically, nature picks s_1, the sender picks m_1, and the receiver picks a_1. They coordinate, and so reinforce their behaviour. Suppose on the second round, choosing stochastically, nature picks s_2, the sender picks m_1, and the receiver picks a_2. They coordinate again, and so reinforce again. Now, message 1 is salient in *both* states for the sender, but no act is salient for any message for the receiver.

[27] The self-correcting nature of the salience game is most apparent when we set $\mu = 1$. In the first round, nothing is salient for anyone, so both the sender and receiver effectively choose randomly. However, if they coordinate, it becomes impossible for them to use salience in different ways because this will lead to miscoordination, and so will never be reinforced.

[28] To see how this might be the case, we might think that a landmark is a naturally salient feature of some geography. We might think such a landmark is *always* salient in the objective interpretation; however, it may be obscured by fog (for example).

If we suppose that the sender and receiver have already started to evolve toward a signaling system—i.e., each state has a unique "most probable" signal and each message has a unique "most probable" act—then a second interpretation of our model is that the salience parameter gives a formal analogue for something akin to *demonstration* in communication contexts. In this case, we might understand the salience game as modeling an interaction of the following sort: an individual sees the state of the world and sends a signal to another individual while subsequently performing the appropriate action for that particular state of the world. In this case, coordination always happens. In the salience game, if the sender and receiver have already started to evolve likely dispositions toward a signaling system, then they will always coordinate when taking advantage of salience. Thus, we interpreted the salience parameter as modeling some degree of salience; however, since the sender and receiver always coordinate when the signal and act are mutually salient, we can interpret this special case as something akin to demonstration, as was just described.[29]

This interpretation is closely related to previous work by Barrett and Skyrms (2017) on a constituent of 'self-assembling games', which they refer to as *cue-reading* games. The cue-reading game is effectively a signaling game where the sender's dispositions are already fixed at the outset. As such, the receiver must evolve to take advantage of this fixed disposition. (Note that under this model we can interpret a sender as the receiver of nature's signal, where nature has a fixed disposition.) However, in the salience game, we do not presuppose that the sender has a fixed disposition at the outset. Rather, the sender evolves a disposition based on precedent—namely, accumulated rewards.

Several questions surrounding social learning arise with respect to salience in nature. In particular, the model we examined here is concerned primarily with how and whether such salience will help effective communication to evolve—and so help individuals in a population avoid sub-optimal pooling strategies, converge toward a convention more quickly, etc.—and if so, to what degree?

We saw that Skyrms' initial motivation in examining signaling from a dynamic learning perspective was to consider the "worst case scenario in which natural salience is absent and signaling systems are purely conventional" (2010, p. 8). However, the models examined here are also ones in which the signaling systems are purely conventional. The salience is not presupposed but evolves naturally as a result of the learning dynamic. The main difference between the atomic signaling game on this learning dynamic and the salience game, then, is that the variant choice function allows individuals to take advantage of their own learning efforts in order to create salience from a purely conventional signal.

---

[29] Note that we stipulate that they evolve up some dispositions at the outset, so that we are not building meaning into the model. The signals are still arbitrary, and so conventional, but in the salience case, when a signal is salient the sender might demonstrate this, and both the sender and receiver coordinate and reinforce.

Further, though salience presupposes little more than the simple reinforcement learning model—only the choice function of the actors is different—it has several desirable properties over and above the atomic signaling game with simple reinforcement learning. Namely, a small amount of salience can help players learn to signal more quickly, and salience can help individuals to avoid sub-optimal equilibria. This is a practical consequence of our results. In particular, this has to do with the evolution of signaling conventions specifically.

A theoretical consequence of this result is that this sort of approach to modeling—taking advantage of parametrization, rather than making assumptions about certain components of the model—allows us to make more realistic assumptions about the phenomena we are trying to model. This parametric methodology thus has consequences for the explanatory value of the model itself. In particular, we have avoided some idealization and obtained results for a continuum of possibilities. Now, we can can say within some interval of a given parameter, what sort of phenomena we would expect to see in a natural setting and how subtle changes to the parameter will affect the outcomes of the game.

References

1. Alston, W. (1964). *Philosophy of Language*. Englewood Cliffs, NJ: Prentice-Hall.
2. Argiento, R., Pemantle, R., Skyrms, B., Volkov, S. (2009). Learning to signal: analysis of a micro-level reinforcement model. *Stochastic Processes and Their Applications*, 119, 373–390.
3. Barrett, J. (2006). Numerical simulations of the Lewis signaling game: learning strategies, pooling equilibria, and evolution of grammar. Technical Report, *Institute for Mathematical Behavioral Science*.
4. Barrett, J. A., Skyrms, B. (2017). Self-assembling games. *The British Journal for the Philosophy of Science*, 68(2), 329-353.
5. Beggs, A. W. (2005). On the convergence of reinforcement learning. *Journal of Economic Theory*, 122, 1–36.
6. Blume, A., DeJong, D., Neumann, G., Savin, N. E. (2002). "Learning and Communication in Sender-Receiver Games: An Econometric Investigation." *Journal of Applied Econometrics* 17, 225–247.
7. Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, 108, 15647–15654.
8. Herrnstein, R. J. (1970). On the law of effect. *Journal of Experimental Analysis of Behavior*, 13, 243-266.
9. Hofbauer, J., Huttegger, S. (2008). The feasibility of communication in binary signaling games. *Journal of Theoretical Biology*, 254, 843–849.

10. Hopkins, E., Posch, M. (2005). Attainability of boundary points under reinforcement learning. *Games and Economic Behavior*, 53, 110–125.
11. Huttegger, S. M. (2007). Evolution and the Explanation of Meaning. *Philosophy of Science*, 74(1), 1-27.
12. Huttegger, S. M., Skyrms, B., Smead, R., Zollman, K. J. S. (2010). Evolutionary dynamics of Lewis signaling games: signaling systems vs. partial pooling. *Synthese*, 172, 177–191.
13. Lewis, D. (1969/2002). *Convention*. Oxford: Blackwell.
14. Mehta, J., Starmer, C., Sugden, R. (1984a). Focal points in pure coordination games: an experimental investigation. *Theory and Decision*, 36, 163–85.
15. Mehta, J., Starmer, C., Sugden, R. (1984b). The nature of salience: an experimental investigation of pure coordination games. *American Economic Review*, 84, 658–73.
16. Pawlowitsch, C. (2008). Why evolution does not always lead to an optimal signaling system. *Games and Economic Behavior*, 63, 203–226.
17. Quine, W. V. O. (1967). Truth by convention. In *Philosophical Essays for Alfred North Whitehead* (pp. 90-124). New York, NY: Russell & Russell.
18. Roth, A., Erev, I. (1995). Learning in extensive form games: experimental data and simple dynamical models in the intermediate term. *Games and Economic Behavior*, 8, 164-212.
19. Russell, B. (1922). *Analysis of mind*. New York, NY: The MacMillan Company.
20. Schelling, T. C. (1960/1980). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
21. Schreiber, S. J. (2001). Urn models, replicator processes, and random genetic drift. *SIAM Journal on Applied Mathematics*, 61, 2148–2167.
22. Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, micro-economics and behavioural ecology. *Current Opinion in Neurobiology*, 14, 139–147.
23. Schultz, W., Dayan, P., Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593–1599.
24. Skyrms, B. (1996/2014). *Evolution of the social contract*. Cambridge: Cambridge University Press.
25. ———. (2004). *The stag hunt and the evolution of social structure*. Cambridge: Cambridge University Press.
26. ———. (2010). *Signals: evolution, learning & information*. Oxford: Oxford University Press.
27. ———. (2010b). The flow of information in signaling games. *Philosophical Studies*, 147, 155-165.
28. Sutton, R. S. & Barto, A. G. (1998). *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.
29. Steinert-Threlkeld, S. (2016). Compositional signaling in a complex world. *Journal of Logic, Language, and Information*, 25, 379–397.
30. Thorndike, E. L. (1905). *The elements of psychology*. Syracuse, NY: The Mason Press.
31. ———. (1911). *Animal intelligence: experimental studies*. New York, NY: The Macmillan Company.
32. ———. (1927). The law of effect. *American Journal of Psychology*, 39, 212–222.