Journal for General Philosophy of Science
Pre-Print Draft. Please Cite Published Version.

Evolutionary Explanations of Simple Communication

Signalling Games & Their Models

Travis LaCroix

Received: 8 March 2019 / Accepted: 8 July 2019

Abstract This paper applies the theoretical criteria laid out by D'Arms et al. (1998) to various aspects of evolutionary models of signalling. The question that D'Arms et al. seek to answer can be formulated as follows: Are the models that we use to explain the phenomena in question conceptually adequate? The conceptual adequacy question relates the formal aspects of the model to those aspects of the natural world that the model is supposed to capture. Moreover, this paper extends the analysis of D'Arms et al. by asking the following additional question: Are the models that we use sufficient to explain the phenomena in question? The sufficiency question ask what formal resources are minimally required in order for the model to get the right results most of the time.

Keywords Signalling Games · Evolutionary Game Theory · Evolutionary Models · Robustness · Modelling

1 Introduction

To communicate meaningfully, members of a linguistic population must cooperate—i.e., they must agree upon some convention: one individual should use a signal to mean one thing if (most) other individuals use that signal in the same way. Conventional language-use in a population of speakers can be modelled using game-theoretic tools; this idea gives rise to the *signalling-game* framework (Lewis, 2002/1969). To explain how simple communication might *evolve*, contemporary scholars—following the initial work of Skyrms

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

(2014/1996, 2010a)—have built increasingly sophisticated models which purport to capture readily observable communicative phenomena in human and non-human populations alike.

Shortly after the publication of Skyrms' (2014/1996) evolutionary account of social norms, and the introduction of his evolutionary explanation of meaning, D'Arms et al. (1998) pointed out how surprisingly little critical attention this use of evolutionary models had received. Their analysis focused on Skyrms' explanation of justice rather than his explanation of meaning. Particularly, they offered three general criteria which an evolutionary model must satisfy in order to pay its explanatory debts: it must be representative, robust, and flexible. Since then, a significant amount of work has been done applying the same evolutionary principles in the context of signalling games and the evolution of language. The models in these various works have become more sophisticated and increasingly subtle; however, the practice of modelling the evolution of communication using the signalling-game framework has still received very little attention with respect the the criteria laid out by D'Arms et al. (1998).

The question that D'Arms et al. (1998) seek to answer can be formulated as follows: Are the models that we use to explain the target phenomena *conceptually adequate*? The conceptual adequacy question relates the formal aspects of the model to those aspects of the natural world that are supposed to be captured by the model. D'Arms et al. (1998) ask the question with respect to evolutionary explanations of *justice*; however, it is not clear that the particulars of their analysis apply, either in degree or in kind, to evolutionary models of *communication*, since this is a different sort of natural phenomenon. This paper applies their critical analysis to various aspects of evolutionary models of signalling.[1]

Moreover, this paper extends the analysis of D'Arms et al. (1998) by asking the following additional question: Are the models that we use *sufficient* to explain the target phenomena? The sufficiency question asks what formal resources are minimally required in order for the model to get the right results most of the time. It should be clear that these questions are closely related in how they inform one another. In order to know *what* the right results are to answer the sufficiency question, the conceptual adequacy question must be appropriately addressed in the first place. A model can be said to be sufficient and conceptually adequate if both of these questions are answered in the affirmative.

Section 2 offers some background regarding the Lewis signalling game and highlights how evolutionary accounts of signalling help to surmount certain explanatory problems that arise from Lewis' initial attempt to model signalling with (classical) game-theoretic tools. Section 3 describes the three initial cri-

[1] There is, of course, a broader concern about modelling and simulations, in general, and their explanatory significance; however, as a first step toward a critical analysis of the ever-increasing literature on signalling games, I will here take a more narrow focus. However, for more general discussion of modelling see, e.g., Sugden (2000) and Humphreys and Imbert (2012).

teria that an evolutionary model must fulfil in order to answer the conceptual adequacy question in the affirmative. Section 3.1 analyses various starting assumptions made when modelling the simplest types of signalling games with respect to the representativeness criterion. Section 3.2 sees whether the computational results fulfil the robustness criterion under varied initial parameters or various extensions of the simplest models. Section 3.3 examines the flexibility of the models with respect to different interpretations of their formal structures. Section 4 examines the sufficiency question with reference to evolutionary models of signalling. Finally, Section 5 summarises the key results of this analysis and suggests how future work might go.

2 Signalling Games

This section offers a brief overview of the signalling game, in addition to some motivations for the move from classical models (Lewis, 2002/1969) to evolutionary models (Skyrms, 2014/1996). This section also introduces some technical language and formal results which will be helpful in subsequent sections.

2.1 Communication Conventions

The signalling game was introduced by Lewis (2002/1969) to provide a formal model for the establishment of conventions. Using the tools of classical game theory, he formalised interactions in which two players use arbitrary signals (messages) to transmit information. This gives a naturalistic account of the emergence of meaning. His key insight is that successful communication requires cooperation.[2]

The signalling game is a coordination game between two players, whom we will call *sender* and *receiver*. For the simplest case, suppose there are two possible states—s_1 and s_2—two possible messages—m_1 and m_2—and two possible actions—a_1 and a_2. We will suppose that there is an appropriate act for each state; in particular, a_i is appropriate in s_i. Call this a 2×2 signalling game, with the dimension of the game referring to the number of state-act pairs by the number of messages. This can be extended in an obvious way to an $n \times n$ signalling game, or, when the number of messages differs from the number of state-act pairs, an $n \times m$ signalling game.

A *strategy profile* in the game is a complete course of action for a player. Table 1 shows all the possible strategies for the sender and receiver in the 2×2 signalling game.

Once we couch communication as success or failure in coordinating (e.g., meanings to state-act pairs), it is necessary to provide some analysis of coordination problems themselves in order to determine how they might be solved.

[2] Cooperation in this sense requires a notion of joint action—the idea that individuals have shared intentions and, perhaps, awareness of their roles. See Gilbert (1989); Cohen and Levesque (1991); Searle (1995); Clark (1996); Bratman (1999).

Table 1: Sender and receiver strategies for the 2×2 signalling game

Sender Strategies		**Receiver Strategies**	
σ_1:	m_1 if s_1, m_2 if s_2	ρ_1:	a_1 if m_1, a_2 if m_2
σ_2:	m_2 if s_1, m_1 if s_2	ρ_2:	a_2 if m_1, a_1 if m_2
σ_3:	m_1 if s_1, m_1 if s_2	ρ_3:	a_1 if m_1, a_1 if m_2
σ_4:	m_2 if s_1, m_2 if s_2	ρ_4:	a_2 if m_1, a_2 if m_2

Coordination problems can be analysed from a game-theoretic perspective, giving rise to equilibria concepts—namely, situations wherein each actor has done the best she can, given what others are doing. The foremost equilibrium concept in classical game theory is the *Nash equilibrium* (Neumann and Morgenstern, 2007/1944; Nash, 1951). This consists of a strategy for each player whereby no player can unilaterally deviate from her strategy in order to increase her payoff. A *strict* Nash equilibrium is one in which a player is strictly worse-off after unilateral deviation from her strategy. In equilibrium, no single player has incentive to deviate from her strategy. This is true even if *multilateral* deviation—several players simultaneously changing strategies—would result in a better outcome for the group as a whole. Table 2 shows the payoff matrix for the 2×2 signalling game. Note that there are two possible ways

Table 2: Payoff table for combinations of strategies in the 2×2 signalling game

Sender \ Receiver	ρ_1	ρ_2	ρ_3	ρ_4
σ_1	1	0	$\frac{1}{2}$	$\frac{1}{2}$
σ_2	0	1	$\frac{1}{2}$	$\frac{1}{2}$
σ_3	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
σ_4	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

in which the sender and receiver can achieve perfect coordination in this case. Lewis referred to these combinations of strategies as *signalling systems*. As such, the 2×2 signalling game has 2 possible signalling systems. These are shown in Figure 1. In general, there are $n!$ possible signalling systems in the $n \times n$ signalling game.[3]

Lewis (2002/1969) points out that conventions may be maintained through prior agreement, but this will not help to quell sceptical worries of how a community comes to agree upon a convention without already having a lan-

[3] Though the symmetric case is not stated explicitly, Lewis (2002/1969) proves that, in an asymmetric signalling game with m states and n signals ($n \geq m$), there are $\frac{n!}{(n-m)!}$ possible signalling systems. As such, when $m = n$, as we have here, it follows immediately that there are $n!$ possible signalling systems.

Fig. 1: The two signalling systems of the 2×2 signalling game

(a) Signalling system 1 (b) Signalling system 2

guage in place (Russell, 1922; Quine, 1967). Even so, Lewis (2002/1969) was able to show that prior agreement is not necessary for convention because the strict Nash equilibria of a pure-coordination game—the signalling systems of the game—are self-enforcing. Furthermore, signalling games have symmetric properties that *entail* the conventionality of language: since every Lewisian signalling game necessarily has more than one signalling system, *which* signalling system a community settles on is entirely arbitrary. What matters is that the community prefers to play one signalling system over the other(s) given that (almost) everyone else in the population prefers to play that signalling system.[4]

However, while symmetry explains the conventionality of language, the sceptical question comes back around as a problem of *symmetry breaking*. It is unclear how a population 'chooses' one signalling convention over another. Lewis suggests that *natural salience*—i.e., a choice-point that naturally stands out from the others—is sufficient for this purpose. Even so, a model that utilises the machinery of classical game theory inherits all of the rational baggage that the theory carries.[5] It is not generally clear how symmetries might be broken spontaneously without either a prior language in place, untenable assumptions about the players' rationality, or potentially *ad hoc* explanations that rely on natural salience. So, it is not clear how the problem of symmetry-breaking can be solved without begging the question.[6]

With this in mind, Skyrms (2014/1996, 2004, 2010a) further developed the simple Lewisian model using *evolutionary* (as opposed to classical) game theory in both learning and biological contexts. An evolutionary model consists of an underlying game and a *dynamic* which determines how individuals' strategies change over time. The idea is that the evolutionary dynamics of the model (whatever they may be) will carry a population to one or another signalling

[4] See Lewis' formal definition of conventions (Lewis, 2002/1969, 78-9).

[5] This 'rational baggage' is exemplified by Lewis' (2002/1969) discussion of *higher-order expectations* in the coordination game: "In order to figure out what you will do by replicating your practical reasoning, I need to figure out what *you* expect *me* to do. I know that, just as I am trying to figure out what you will do by replicating your reasoning, so you may be trying to figure out what I will do by replicating 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" (27-28). Without an excessive requirement on the rationality of the players, or a prior language already in place, it is not clear how Lewis' notion of signalling can arise spontaneously.

[6] See Skyrms (1990, 2014/1996); Vanderschraaf (1995).

system as the game is repeated. Thus, evolutionary models give a non-question-begging explanation of how symmetries are broken without reference to natural salience, rationality, or prior (or tacit) agreement: the dynamics of the system itself breaks symmetries randomly or dependent upon the starting point of the system.

Two related questions arise in light of evolutionary models: Can the meaning of a signal emerge spontaneously, by chance? Can signals spontaneously acquire information through naïve learning in repeated interactions? Evolutionary models of signalling illustrate how and when a meaningful term-language might evolve from initially meaningless random signals by way of adaptive fitness and evolutionary dynamics. A Darwinian model of evolution by differential reproduction, natural variation, and mutation gives an affirmative answer to the first question, at least given certain starting assumptions. This is a process of biological, or phenotypic, evolution of strategies. Trial-and-error learning gives a way to model cultural, or economic, evolution of strategies. This gives an affirmative answer to the second question, given certain starting assumptions. As it turns out, these are two different ways of saying essentially the same thing: the two models are closely related mathematically. Once a simple signalling language is up and running, we can ask the further questions: How much information does an individual signal carry? What is the informational content of an individual signal? These questions can be answered with more technically complex models which utilise information theory.[7] Out of this comes a highly interdisciplinary family of models that purport to explain how language, meaning, and information transfer may have arisen through evolutionary processes given minimal starting assumptions.

Nonetheless, while evolutionary models surmount the problems to which classical models give rise, the framework and starting assumptions for these evolutionary models have received little critical attention. We turn now to the *conceptual adequacy* and *sufficiency* questions raised above.

3 The Adequacy of Our Models

This section details the criteria offered by D'Arms et al. (1998) for assessing the conceptual adequacy of evolutionary models and offers motivation for asking the conceptual adequacy question in the first place. Once the terminology is described, each criterion is applied to various aspects of the evolutionary models with respect to signalling games.

Mathematical models are used in a variety of disciplines with several goals in mind. Models in the natural sciences accomplish these goals with two primary interrelated methods: (1) they abstract away the irrelevant features of real-world phenomena while retaining those features that are (presumed) relevant, and (2) they make simplifying assumptions in order to grant some degree of mathematical tractability to the modeller. There is obviously some theoretical overlap between these two points, but the main challenge is that the

[7] In particular, Kullback-Leibler divergence. See Skyrms (2010a,b).

modeller must be particularly sensitive to (1) exactly what is relevant and what is not, and (2) what the acceptable trade-off is between computational simplicity and accurate results.

Batterman (2009) points out that idealising is "a means for focusing on exactly those features that are constitutive of a [phenomenal] regularity" (430), which is to say the *essential* features of a particular phenomenon. Note, however, that understanding exactly *what* the essential features of phenomena are is paramount in building a model. In particular, it is important that the data obtained from the model be interpreted appropriately, lest one interpret artefacts that arise from the mathematical structure, as opposed to the phenomena the model is supposed to represent, as significant.

D'Arms et al. (1998) give three basic criteria for assessing the value of evolutionary models in solving philosophical problems. The specific models they examine concern the evolution of *justice*. Here, we are concerned with the evolution of communication. Nonetheless, the meta-theoretical criteria that D'Arms et al. (1998) offer should still apply. In general, if it is to be conceptually adequate, an evolutionary model should be (1) representative, (2) robust, and (3) flexible.

For a model to be representative of a phenomenon, it must be the case that "[c]ircumstances with the structure of the mathematically characterized interaction which the model treats must be realized with sufficient frequency in the environment of evolutionary adaptation" (D'Arms et al., 1998, 89). D'Arms et al. (1998) characterise two distinct senses of *robustness*. On the one hand, a model is robust when it gives the desired result across a variety of parameters—this is sometimes called *structural stability*; with respect to a particular model, it is said to be robust when the results it gives are stable under perturbation of starting conditions (89). Flexibility is partitioned into two conditions. First, the evolutionary strategy's adaptiveness that the model supposedly models can be realised by several different mechanisms, and, second, the model itself can be interpreted as representing different possible processes (89).

D'Arms et al. (1998) show how Skyrms' (2014/1996) evolutionary account of justice succeeds in several different ways on the robustness and flexibility criteria but fails on the representativeness criterion. Further, they show that as the representativeness of the model increases, the flexibility and robustness of the model actually decrease. However, their original criticisms do not obviously apply to evolutionary models of *signalling* behaviour. For example, one charge against Skyrms (2014/1996), with respect to representativeness, is that there is no real-world analogue of, e.g., the 'referee' component in the underlying game he uses to explain the evolution of justice;[8] however, the signalling game has no such component, so this criticism explicitly does not apply here. Sections 3.1, 3.2, and 3.3 address each of the model-theoretic criteria in turn and use these criteria as a means to assess the adequacy of the signalling-game framework in general.

[8] See Skyrms (2014/1996) for details of this model.

There is an important assumption to make clear in considering the criticism of D'Arms et al. (1998) in light of the evolutionary signalling models. Their analysis was solely concerned with evolutionary explanations of *justice*. In this case, the *target phenomena* (to which I referred in the introduction) to be explained is justice, which is cashed out in terms of *fairness*—specifically, a fair split in a Nash demand game. It seems obvious that forwarding a medium demand to one's opponent in this framework captures a clear and intuitive notion of 'fair'. However, the Nash demand game and the signalling game are not analogous in this sense: it is not entirely obvious *what* exactly the explanatory target in the signalling game is supposed to be.

A number of related questions arise here. For example, do we want to explain how signalling arises in the simplest case—the 2×2 Lewis signalling game? Do we want to explain how we can avoid partial-pooling equilibria in more complex games? Do we want to show how efficient communication conventions arise when there are more states/actions than messages? And so on. The ambiguity of what our explanatory target for signalling games actually consists in is further complicated by the fact that communication often fails in real life.[9]

Ultimately, the goal of any research in language origins is to explain how possibly *language* evolved. An obvious, and oft-mentioned difficulty that arises in such research is that we have no direct evidence to observe in support of any given theory—for example, whether complex syntax itself was subject to selective pressure (Pinker and Bloom, 1990; Jackendoff, 1999, 2002; Pinker and Jackendoff, 2005) or sexual selection (Progovac, 2015), whether syntax was a byproduct of physical changes that were themselves subject to selective pressure (Lieberman, 2000), or whether it was purely the result of cultural evolution (Tomasello, 1999). Nor can we go back in time to see what actual-world precursors were in place to drive the evolution of this complex system.

In lieu of a time machine, analytic and simulation results can "provide a novel empirical way of testing plausibility of evolutionary hypotheses, even when they cannot themselves directly confirm or refute such hypotheses" (Progovac, 2019, 61). In order to obtain such results, however, we first need a model, and evolutionary game-theoretic models can shed light on a variety of linguistic phenomena—at least in terms of a *proof of possibility*. Of course, many researchers who utilise formal models acknowledge the limitations of these tools, but they are helpful in conjunction with empirical evidence from, e.g., evolutionary biology and linguistics. Thus, signalling games can be understood as a useful tool in the sort of multi-component approach that is advocated by Fitch (2010, 2017).

That being said, the type of communication system that is well-modelled by the simplest signalling game is extremely simple. Even the oft-cited case of

[9] Note, however that this is also true of *fairness* as a target phenomenon to be explained by an evolutionary model.

the alarm-call signalling systems of vervet monkeys are not, strictly speaking, signalling systems in the Lewisian sense.[10]

Even so, prior to Skyrms' evolutionary analysis of the spontaneous emergence of meaning, it was a genuine puzzle how this might happen. Skyrms (2010a) suggests, however, that "[t]here is no mystery behind the emergence of signaling" (177), since, under the right circumstances, adaptive processes can lead to the spontaneous emergence of meaningful signals. This is a small, but extremely important, piece of the larger puzzle of how bona fide languages emerge. While the signalling-game framework does not (as of yet) provide a solution to this problem, the simple evolutionary models suggested by Skyrms provide a foundation for future work in language origins research to build upon; as such, it is important that the foundation is deemed to be solid.

3.1 Representativeness

The representativeness of evolutionary models with respect to signalling is, perhaps, the most difficult criterion to be assessed. The phenomenon to be explained is that language exists; the pertinent questions is "How could language have arisen in the first place?" Signalling games appear to give a *prima facie* answer that a community's evolution toward some kind of proto-language is at least possible—this is certainly a necessary condition for the existence of communication—and they also appear to explain *how* it is possible.

However, though formal models must represent phenomena to some degree in order to be effective, there has been relatively little discussion in the philosophical literature about what precisely this means with respect to signalling games. One exception is Sterelny (2012), who is diffident about whether Skyrms' evolutionary explanation of signalling really is informative. This is because the evolutionary signalling game "abstracts entirely away from proximate mechanism" and 'black-boxes' the organism, thus betting on "the independence of historical trajectories from mechanical implementation" (84-85).[11] However, as was mentioned in the previous section, the 'first-step' phenomenon to be explained is merely the emergence of meaning, in order to show that this is possible without deference to, e.g., salience or prior agreement.[12]

Skyrms initially showed, using numerical simulations, that in the 2×2 signalling game, something like meaning arises spontaneously as a "moral cer-

[10] Sterelny (2012) highlights that the underlying problem that the vervet alarm-call system solves is decidedly *not* a coordination problem, since a receiver vervet's payoff for, e.g., running up a tree (upon hearing bark) does not in any way depend upon the sender vervet's running up a tree. Thus, the vervets do not have mutually dependent rewards (76).

[11] Proximate explanations are contrasted with (and complementary to) ultimate explanations. The latter would require detailing the evolutionary trajectory by which signalling arises in addition to the selective forces driving dynamic changes in the population, whereas the former would require detailing the developmental and physiological mechanisms by which signalling is implemented. See Mayr (1961).

[12] See also Brusse and Bruner (2017) for an explicit response to several of the worries brought up in Sterelny (2012).

tainty".[13] Huttegger (2007a) gives an analytic proof that, in the 2×2 signalling game, infinite populations always converge to a signalling system with probability 1 under the replicator dynamic.[14] Pawlowitsch (2007) finds the same result for finite populations, assuming the frequency-dependent Moran process.[15] Similarly, Argiento et al. (2009) prove convergence to a signalling system in the limit under a simple reinforcement learning dynamic.[16] Thus, given certain starting assumptions, senders and receivers evolve toward one or the other signalling system—a state of perfect (or near perfect) communication—in the simplest case—i.e., the 2×2 signalling game with unbiased nature.

However, even in the simplest case the assumptions that the model utilises in order to obtain such promising initial results are questionably representative.[17] Of course, some flexibility should be granted to the representativeness of our assumptions as long as there is adequate justification for making assumptions that are unrepresentative. It is commonplace for, e.g., models of evolutionary biology to assume that populations are infinite or continuous because these simplifying assumptions allow such models to inherit the power of the calculus. So, the results more than pay for the cost of the assumptions. If representativeness is to be taken seriously, then each of the assumptions made in the model needs to cash out at the 'accurate representation' check-out counter *or* the 'payoff of results' check-out counter.[18]

Some of the assumptions that go into even the simplest signalling game models include, but are not necessarily limited to, (1) population size, (2) payoff structure and interpretation, (3) whether or not pairing is random, (4) whether or not sender/receiver strategies are initially random, (5) whether or not states are equiprobable. I offer a brief analysis (and justification, where appropriate) of each of these assumptions in turn.

Infinite Populations. The assumption that a population is infinite is perhaps the most obviously unrepresentative of all the assumptions made in signalling

[13] Here, Skyrms is referring to a definition in Jacob Bernoulli's *Art of Conjecture* where he says that something is morally certain if its probability comes so close to complete certainty that the difference cannot be perceived. However, it should be noted that in light of subsequent work in this area, Skyrms weakened this claim in a later edition of the book. The claim itself is still true in a particular case, it is just not true generally.

[14] The replicator equations capture the idea that an individual with higher-than-average fitness is more likely to reproduce; the composition of the population changes over time, resulting in corresponding changes to the fitness of a particular strategy relative to the average fitness of the population. See Taylor and Jonker (1978) for further details.

[15] In this case, an individual in the (finite) population is selected proportional to its fitness; that individual produces an identical offspring which then replaces a *randomly chosen* individual in the population. Crucially, it is possible for a single mutant strategy that has a disadvantage with respect to relative fitness to generate a lineage that eventually takes over the entire population. See Nowak et al. (2004); Nowak (2006) for further details.

[16] Herrnstein reinforcement learning, based on the *matching law* (which, in turn, is a formalisation of the *law of effect*), supposes that the probability of selecting a particular action is proportional to the accumulated rewards for that action. See Thorndike (1905, 1911, 1927); Herrnstein (1970) for further details.

[17] More complex cases will be examined in Section 3.2.

[18] To borrow a turn of phrase from John Woods.

game models. Nonetheless, assuming that a population is able to become arbitrarily large can ease computational complexity, as limits are allowed to tend toward infinity. However, the concept of a limit itself may be conceptually suspect insofar as signalling conventions must be arrived at in finite time in order for the model to explain how signalling may have actually arisen in nature. Still, no one claims that this is in fact representative of the actual-world conditions wherein signalling arises. So, it seems that this assumption cannot be cashed out at the 'accurate representation' check out counter, even if it gets the long-term qualitative behaviour quite right.

It should be noted that, in an evolutionary model, populations can refer to strategies rather than individuals *per se.* As such, the concept of an 'infinite population' might be interpreted as the existence of an infinite number of strategies—e.g., when individuals mix over their strategies with some probability distribution. Even so, if we assume that there is an infinite number of strategies, then one of two things must follow: there is an infinite population of individuals each playing a finite (or infinite) number of strategies, or there is a finite population of individuals each playing an infinite number of possible strategies. In either case, it is not clear that this is going to be representative.

Nontheless, for the 2×2 game, Pawlowitsch (2007) shows that the same results for infinite populations can occur assuming finite populations, so infinite populations are not *necessary* for the desired results. The utility of assuming infinite populations, then, obtains precisely in the simplicity of the mathematics being used on them and the tractability of questions about robustness under perturbations. So, even though this assumption is unrepresentative, it pays for itself with the computational power it affords to the model. As such, the payoff of relative computational simplicity pays for the unrepresentativeness of the assumption.

The assumption of infinite populations is closely related to whether populations should be understood as continuous or discrete. If we denote the proportion of a population playing a given strategy S_i of the entire strategy set S at a time t_0 as $P(S_{i_{t_0}})$, then the discrete update of proportions for the population at the next time, t_1, is given by $P(S_{i_{t_1}}) = P(S_{i_{t_0}}) \left[\frac{f(S_i)}{f(S)}\right]$, where $f(S)$ is the average fitness of all the strategies in the population, and $f(S_i)$ the fitness of a particular strategy i.

If a strategy S_i has high fitness, then $f(S_i)/f(S) > 0.5$, and so a larger proportion of the population will play strategy S_i at time t_{n+1} than was the case at time t_n. Similarly, if S_i has low fitness, then $f(S_i)/f(S) < 0.5$, and so the proportion of players playing that strategy at a future time, t_{n+1}, will decrease. Certainly updates of strategies will more naturally be associated with, perhaps irregular, discrete time intervals. Nonetheless, the continuous-time model of the dynamics of a signalling game is simply an idealisation of the discrete-time model: namely, the rate of change of a strategy in a population is given by *differential reproduction*,

$$\frac{dP(S_i)}{dt} = P(S_i)\left[f(S_i) - f(S)\right].$$

This *replicator equation*, due to Taylor and Jonker (1978), is essentially the application of a limiting process to the discrete-time dynamics.[19] If populations are finite, then analysis of the discrete-time dynamics is actually quite difficult. As such, there is a computational payoff for assuming some sort of continuous dynamics for finite or infinite populations. Given the (relative) computational simplicity in analyses of differential equations, or systems of differential equations, this assumption cashes out at the 'payoff of results' check-out counter.

Payoff Structure and Interpretation. There are several interpretations that one could make of the 'payoff' component of the signalling game. Unlike a lab setting, subjects in a population may not literally be rewarded (punished) when they achieve (fail to achieve) coordination. Some interpretations of payoff may presuppose too much of an intentionality component. In Lewis' (2002/1969) initial formulation, he gives an example of a meeting as a coordination problem: if both players *prefer* to meet, then coordinating will pay the same for each player when they coordinate and when they fail to coordinate. However, this is not necessary. It is also perfectly reasonable to interpret 'payoff' less literally. In this case the word *payoff* is just a name for some evolutionary structure—we call it *payoff* as a linguistic convention for the ease of reference and mathematical computation. But the idea of a payoff is not necessarily imposed by the technical components of the model itself; rather, payoff simply adds "a fitness component to a birth-and-death process that introduces an element of frequency-dependent selection in addition to drift" (Pawlowitsch, 2007, 612). Indeed, the payoffs are arbitrary in some sense, and signalling games with a variety of payoff structures have been examined. In this case, the particular payoff structure that the modeller chooses to use will need to be justified on a case-by-case basis.

The interpretation of payoff structure is closely related to the assumption that signalling games lie arbitrarily close to the *pure* coordination end of the spectrum (Lewis, 2002/1969).[20] The assumption that signalling games are games of pure coordination is not necessarily representative. The examples that Lewis (2002/1969) gives are ones in which the players obviously desire the same outcome—e.g., meeting one another somewhere, calling back or waiting if a telephone call is disconnected, etc. However, real-world considerations call this assumption into question. For example, in alarm calls of vervet monkeys (Cheney and Seyfarth, 1990), sending a signal and possibly alerting a predator to the sender's whereabouts makes the payoff apparently higher for those players that were initially unaware that a predator was nearby: the choice to not send a signal gives high payoff to the potential sender (who has seen the predator), and lower payoff to the receiver(s). Thus, the players' interests are not *purely* coordinated.

[19] See also, Zeeman (1980); Hofbauer and Sigmund (1998, 2003); Skyrms (2009, 2010a).

[20] That is, each player gets exactly the same payoff. The spectrum that Lewis (2002/1969) refers to is due to Schelling (1980/1960), with games of pure coordination at one extreme, and games of pure conflict (i.e., zero-sum games) at the other extreme.

If pure coordination is not representative of the actual-world phenomena, then this may be problematic for the model. It is often assumed that in a zero-sum game no meaningful signalling could arise precisely because the players' interests are purely opposed. As such, it will always be in the one player's interest to send *the wrong* signal given the state. As Franke et al. (2012) point out: it "is easy to see that under conditions of extreme conflict (a zero-sum game), no informative communication can be sustained. For why should we give information to the enemy, or believe what the enemy tells us" (26). However, Wagner (2009) shows that meaning can arise even in a zero-sum game: the dynamics of the system is chaotic, but in spite of the lack of equilibria, the signals used in the game are still meaningful—i.e., they carry at least some information. Franke and Wagner (2014) ask how meaning can evolve when preferences are not equivalent. They point out that with the exception of the Sir Philip Sydney Game (Maynard Smith, 1991) and Spence's (1973) Job Market Signalling Game, the question is still largely open.[21]

Random Pairing. In constructing a dynamical model, a decision must be made as to how and when individuals are chosen to interact. Skyrms (2014/1996) initially assumes that pairing is random in his evolutionary account of justice, and the same model is built for an evolutionary account of signalling. When pairing is not random, there may be a correlation mechanism at play. Indeed, this is a point which D'Arms et al. (1998) criticise in the evolutionary explanation of justice, and their criticism applies equally well to an evolutionary explanation of communication. A correlation mechanism is a parameter, ϵ. Here $\epsilon = 0$ means that pairing is random, and $\epsilon = 1$ means that pairing is perfectly correlated. Once a correlation mechanism is introduced into our model, the question might be asked where in the interval $[0, 1]$ ϵ should lie in order to achieve some degree of representativeness.

However, D'Arms et al. (1998) note that it is equally possible to assume an *anti*-correlation mechanism. This is shown to have significant effects on otherwise robust results. Both of these mechanisms are discussed in Skyrms (2000). In the context of Skyrms (2014/1996), the correlation parameter is a *global* parameter. However, this too is an idealisation: interactions in any given real-world population may be correlated in some ways and anti-correlated in others. Thus, as far as representativeness is concerned, it would be a mistake to assume that real populations can be completely accurately represented by a single correlation parameter, even if many populations may be approximated in this way.

The signalling game does not need to be played in abstract space—the model can also impose some sort of geography on the population, which will

[21] Note that this question is also taken up in Crawford and Sobel (1982); their results suggest that "perfect communication is not to be expected in general unless agents' interests completely coincide" (1450); however, Ahern and Clark (2014) show that when misalignment of preferences is not too strong, a 'cyclic' signalling system can evolve—they note a "range of behavior, from separating, to cycling, to collapse" as conflict increases (31–32). See also Godfrey-Smith and Martínez (2013); Martínez and Godfrey-Smith (2016).

restrict the types of interactions that may take place. Zollman (2005) examines learning to signal with neighbours on a grid, using imitation dynamics, where each individual is able to observe each of her eight neighbours, and imitates the best strategy that she sees. (The topology of the network is a 100×100 grid, placed on a torus, so that each of the $10,000$ individuals on the grid has eight neighbours.) In this case, it is found that alternative signalling systems are able to co-exist, but occupying different regions of the topography. Wagner (2009) extends this analysis and compares neighbourhood interactions with more complex network structures.

Wagner (2009) shows that the behaviour of the system—i.e., whether and when populations converge to signalling systems under a variety of starting assumptions—depends significantly on the topological structure of the network itself. One of the main results of his argument is that the topological structure of so-called 'small world networks' is very conducive to the efficient evolution of meaning.[22] This is significant because many real-world social interactions *in fact* take place in small world networks.[23] Mühlenbernd (2011) Also examines the evolution of signalling in a structured spatial society. This is similar to the torus-grid networks examined in Zollman (2005); Wagner (2009), except the agents can 'choose' to interact with more distant neighbours in the community. The choice is established by a degree-of-locality parameter, whereby an individual chooses to interact with a neighbour with some probability, determined by the (Manhattan) distance of that neighbour from herself and the degree of locality. This parameter fills the gap between Zollman (2005) and Wagner (2009), whose models are at the extremes of the scale that Mühlenbernd (2011) introduces. (For example, the higher the degree of locality, the more probable that an agent will choose to interact with her immediate neighbours.)

Given that real-world interactions take place in physical space, these extensions capture a notion of correlation that is indeed (more) representative. Social animals, like humans, live in groups and so are more likely to pair with the same individuals repeatedly. Further, constraints on mobility give practical reasons to assume that certain pairings amongst the entire population simply could not have happened. It seems the best we can do in this case is offer a parameter sweep in order to see how small changes to correlation and network structure might affect the outcomes of the signalling game. Assumptions leading to positive or negative results will need to be cashed out as representative or not on a case-by-case basis. It is of some import, however, that correlation generally tilts the balance *in favour* of signalling systems, since net-

[22] *Small world network* is a technical term characterised by a graph with a certain set of properties—e.g., high clustering coefficient (of nodes), short average path length (between nodes), etc. For example, many forms of the underlying architecture of the internet are small-world networks.

[23] Mühlenbernd and Franke (2014) give a nice overview of how different network topologies shift the basin of attraction for signalling systems and pooling equilibria (by assuming non-equiprobable states).

work interactions—as well as kin selection, partner choice, group selection—are correlation-generating mechanisms.[24]

Initially Random Signals/Acts. Similar to the assumption of random pairing between players, a question arises about the representativeness of initially random signals or acts. In the initial formulation in Lewis (2002/1969), it was assumed that natural salience would be one means of arriving at a signalling system. However, while there may have been some natural salience in play in the evolution of signalling, it seems premature to assume what that natural salience may have been. Thus, the assumption in Skyrms (2010a) that there are no initial saliences is a way of showing how signalling can arise even in the hardest possible case, where everything is symmetric.

While it may be the case that natural salience makes it so that signalling does not need to be initially random, if the population can converge to a signalling system with initially random signals or acts, then something like natural salience will help the system to converge more quickly. Indeed, this is demonstrated in LaCroix (2018). As such, we assume less in order to obtain stronger results. While a lack of salience is not necessarily representative, any degree of salience is going to make effective signalling more likely.

Equiprobable States. A final assumption that we will consider is that the states that obtain in the model do so with equal probability—If $N(s)$ is the number of states in a signalling game, and s_i is a particular state, then the probability that s_i obtains is $P(s_i) = 1/N(s)$. It is not clear that an assumption of equiprobable states, as in the simplest model, is representative: one state may obtain more often than another. In the 2-state model, if state one occurs with $P(s_1) = 0.99$, and state two occurs with $P(s_2) = 0.01$, then a signal contains little to no information in this game: the population ends up in a pooling equilibrium where the sender sends the same signal regardless of the state, and the receiver does the same act regardless of the signal. Despite the fact that the signal carries no information about the state, players actually do quite well (Skyrms, 2010a, 64-66). As such, this situation fails to capture the results that we want from an explanation of how *meaning* (i.e., of signals) emerges. Furthermore, the larger the disparity between the state probabilities, the more likely it will be that the population converges on a pooling equilibrium (Skyrms, 2010a).

Nonetheless, under the *replicator-mutator* dynamics, the line of pooling equilibria in the game with significantly unequal probabilities of states collapses to a single point. When the disparity between probabilities is not too great, the pooling point is dynamically unstable (otherwise it is an attractor). Hofbauer and Huttegger (2008) show that this bifurcation occurs around $P(s_i) \in (0.78, 0.79), P(s_j) = 1 - P(s_i)$. As such, even if one state occurs three times as often as the other, populations converge toward signalling systems assuming some mutation in the dynamics.

[24] See also the discussion of correlation in, e.g., Skyrms (1994); D'Arms (1996, 2000); Kitcher (1999); Gintis (2000); Harms (2000); Alexander (2007).

In learning contexts, the results are similar. Skyrms shows that initial weights have a significant effect on whether or not senders and receivers learn to effectively communicate.[25] Intuitively, one might think that a large disparity in state probabilities is less problematic if we assume that the stakes are high: as Skyrms (2010a) points out, "[p]redators may be rare, but it does not pay to disregard them" (67). This is precisely what Hofbauer and Huttegger (2008) show. Their payoff parameter measures the *importance* of communication; even when there is a large disparity between probabilities of states, as long as the information being communicated is important enough, populations can still evolve toward effective communication.

As such, the simplest models seem to capture intuitively representative assumptions about how and when states of the world obtain and how the importance of communicating information, even about improbable states, affects a population's ability to converge toward a signalling system or to learn how to signal.

The considerations above primarily have to do with assumptions made in the simplest models of signalling games. However, more fruitful results are obtained from more complex models. Nonetheless, several of the more complex models make the same or similar assumptions at the outset, and simply extend those assumptions with different parameters—i.e., having more players interacting, having more states and signals, etc. Given the results that have obtained in the literature so far, these more readily fall under the *robustness* criterion. For the most part, many of the assumptions made in creating a model of simple signalling games within a population are *not* representative of the circumstances of the actual world, *per se*; the claim here is that many of the falsities pay their rent by easing computational complexity or by strengthening the results obtained. So, while representativeness is difficult to analyse, several justifications can still be given for the starting assumptions made in the models.

3.2 Robustness

It was previously said that the simplest signalling games see populations converging on signalling systems with near certainty given certain starting assumptions made in the model, and furthermore that several of these starting assumptions can be cashed out as representative—or, if they cannot, then they can at least be cashed out in terms of mathematical tractability. The next criterion to examine is whether or not these results are robust. Recall that for a model to be considered robust, in the sense of structural stability, it should give the desired result across a variety of parameters. For a particular model,

[25] The results of the Roth-Erev model (Roth and Erev, 1995; Erev and Roth, 1998) are quite similar to the results of the Bush-Mosteller reinforcement learning model (Bush and Mosteller, 1955), where learning parameters play a role analogous to initial weights. See the discussion in Skyrms (2010a, 97-98).

robustness requires that it be stable under perturbation of starting conditions. Since the simple signalling game model presented in the previous section is indeed robust in both senses of robustness, we will consider robustness in the sense of structural stability. There are several ways in which the parameters of the model may be varied. The question then is whether the results obtained are the same or similar. We turn now to other extensions of the simplest models to see how and when the results differ.

One natural way to change the parameters of the simple model is to assume that the number of initial states, acts, and signals is greater than 2, resulting in an $n \times n$ game. Another way is to assume that there are either more states and acts than signals, or more signals than states and acts, resulting in an $n \times m$ game. Several of these cases have been analysed via simulation.

To start, let us look at the $n \times n$ game with $n > 2$. The question is whether we get the same results as we did for the 2×2 model. It turns out that we do not. Barrett (2006) extends simple signalling game using an adjustable reference point with truncation (ARP) learning model, due to Bereby-Meyer and Erev (1998). He examines the results of running simulations with $n = 3, 4$, and 8. The results of his simulations are shown in Table 3. Here we

Table 3: Imperfect communication in symmetric signalling games with more than 2 state/act pairs and messages

Model	**Run Failure Rate**
3-State/3-Term	0.096
4-State/4-Term	0.219
8-State/8-Term	0.594

see partial pooling and miscommunication arising very quickly. Nonetheless, Barrett points out that in every case the players coordinate better than chance: information is still transmitted. One should note that Barrett's simulations involve 10^3 runs, each with 10^6 plays. For each run, a signalling game is said to 'converge' to a signalling system if it has a success rate of at least 0.8—where the success rate is calculated as the number of successful plays divided by total number of plays. The 8-state/term case is more open to interpretation as the data is presented here because though more than half of the runs failed to converge to a signalling system with cutoff 0.8, when the cutoff is relaxed to 0.75, only 4.6% of the runs failed.

However, this reiterates the question of what it means for a model to be adequate or sufficient in representing the phenomena in question. We could see robustness in a population's convergence toward a signalling system in almost any case, assuming we relax certain conditions—i.e., the cases in Table 3 could all be shown to converge with near-certainty if we relaxed the parameters to a success rate of 0.5 or 0.2. However, this would be unsatisfying because there is *clearly* no justification for setting the success rate at anything less than chance (in the very least). But if communication evolves at rates better than chance, then what is the justification for picking a cutoff of 0.8 rather than 0.75?[26]

One might think that, even with the more stringent success rate parameters, it is possible that the population *does* converge in the limit—since these are only numerical simulations, perhaps the success would be higher if there were 10^7 or 10^{10} plays in each run. Indeed, this might help to some extent, however this turns out to be somewhat optimistic: while there is no analytic proof of when signalling systems evolve in these more complex games, Pawlowitsch (2008) gives a complete characterisation of partial pooling equilibria, and shows (assuming the replicator dynamics) that the partial pooling equilibria in fact have basins of attraction with positive measure. So, Barrett's negative results are robust.

However, does this mean that the signalling model fails the robustness criterion? This would be hasty. It is true that minor changes in the starting assumptions can significantly change the results, but this is not true of many of the broad results—e.g., whether or not significant information transfer emerges spontaneously. Thus, if our explanation requires that perfect communication must evolve, then the signalling game model will fail the robustness criterion. However, this does not preclude the reliable evolution of some, or even a substantial amount of, communication. Just because perfect communication does not always evolve, this does not undermine the robustness of these models for explaining the conventional origins of communication *qua* information transfer.

Indeed, we can afford to say something stronger than this. Even if almost half of the experiments of 8×8 games do not result in a perfect signalling system under basic conditions, as in Barrett (2006), they result mostly in near-perfect signalling systems. It is true that there exists some partial pooling, but the remaining states/signals/actions have one-to-one mappings. Thus, we can argue that pure signalling evolves robustly, even in these complex cases. The robustness criterion fails only under the strictest conditions for the expected outcome.

It seems fair to relax our expectations here, given that no human language constitutes a perfect signalling system due to the existence of synonyms and polysemy. Thus, the supposed failures of robustness actually capture something representative of human languages. While signalling systems are *global* maxima of the fitness landscape of information transfer, partial-pooling equilibria constitute *local* maxima. The type of (learning or evolutionary) dynamics under consideration determine the probability that a signalling strategy will reach a local or global maximum. This probability can be increased significantly by changing the dynamics to make it 'explore' more—for example,

[26] One such justification is got by analysing the payoffs or success rates for suboptimal pooling strategies. For example, in the 4×4 signalling game, the most efficient pooling strategy has an expected payoff of 0.75, and a success rate of 0.75. A cutoff of 0.80 in this case is justified, since a suboptimal random walk may spend some time above the 0.75 success rate before settling in to a partial-pooling equilibrium. However, the most efficient pooling strategy in an 8×8 signalling game has an expected payoff (and success rate) of 0.875. As such, a cutoff of 0.80 is not warranted here, since this will include many runs that ended up with suboptimal conventions.

adding mutation to the replicator dynamic (Huttegger et al., 2010), adding mutation to the learning dynamic (Barrett, 2006), adding a forgetting parameter to the learning dynamic (Barrett and Zollman, 2009), or allowing for the spontaneous invention of new signals (Alexander et al., 2012; Mühlenbernd and Nick, 2014). Thus, if we require that the trajectories of our dynamics end up in a global maximum in order to be called successful, then the model will certainly fail the robustness criterion; however, if we relax the conditions for success to include near-perfect signalling systems, this will dramatically increase the robustness of the model, since these local maxima can be arrived at under a variety of different conditions.

Another possibility is that there are more than two states, acts, or signals, but they are unequal. In the case where there are more signals than state-act pairs, we get a system with synonyms, and in the case where there are more state-act pairs than there are signals, we get informational bottlenecks. Barrett (2006) also examines these possibilities with the ARP model.[27] He shows that having more terms helps the population evolve toward perfect communication, though there will be informational redundancy. He also shows that in signalling games where the states are not homogeneously distributed, partial pooling appears with greater frequency. It has also been shown that, in the synonym case, no equilibrium is evolutionarily stable (Donaldson et al., 2007). Nonetheless, a correlation mechanism (in the evolutionary dynamics), or some degree of negative reinforcement (in the learning dynamics) can help to decrease the basins of attraction for sub-optimal polymorphic traps.

One final thing to highlight is the close connection between robustness and representativeness, with respect to our target phenomena. If the target, based on the results of the simple 2×2 signalling game, is to show how signalling systems arise systematically, then the results discussed here show that the signalling-game framework is not robust in the sense of structural stability. However, there is no compelling reason to posit such a strict explanatory target. On the one hand, it has been shown empirically that individuals do not come to a perfect signalling system in more complicated cases (Bruner et al., 2018). On the other hand, signalling phenomena in nature is often significantly more complex than the signalling model suggests, so there is not necessarily any reason to think that partial pooling constitutes a 'failure' in any sense of the word.[28]

All of these models are obviously idealised; as such, they are simplifications.[29] However, one aspect of the robustness (i.e., of a model) is that several independent constructions might arrive at the same, or similar, qualitative phenomena. This is precisely what we see in the signalling game. Rather than

[27] See also, Skyrms (2010a).

[28] Both Sterelny (2012) and Santana (2014) highlight that signalling in nature is often one-to-many or many-to-many, rather than the neat and tidy bijection that obtains in a signalling system. Thus, in targeting more complex signalling phenomena than merely the emergence of meaning, pooling should be expected.

[29] For an extensive and systematic discussion of the problem of determining whether the right simplifications have been chosen, see Wimsatt (2007).

being a burden of complexity, each novel model serves as a data point for the robustness of the overall framework of the signalling game—i.e., the simple, core structure upon which such extensions are built. This is the core of Levins' (1966) notion of *robustness analysis*. When a number of similar or distinct models lead to similar results, this can only serve to reinforce the usefulness and strength of the models: "our truth is the intersection of independent lies" (20).[30]

There are, of course, many other ways of extending the simplest signalling game, but this serves to highlight an important feature of signalling games: their models require a significant number of assumptions, and minor changes in the starting assumptions can significantly change the results. Even so, as we have seen, these results are robust on all but the most stringent conceptions of what constitutes 'success'.

3.3 Flexibility

We have seen that models for signalling games perform reasonably well under the representativeness criterion, and reasonably well on the robustness criterion, under some relaxed criteria for success. It remains to be seen how well they do on the flexibility criterion. Recall that there are two conditions to be met here: (1) the evolutionary strategy's adaptiveness that the model supposedly models can be realised by several different mechanisms, and (2) the model itself can be interpreted as representing different possible processes. In fact, signalling game models perform quite well on the flexibility criterion.

On the one hand, the actual-world adaptiveness of the signalling system strategy can be realised through biological processes, which are modelled by the replicator dynamic. This constitutes a model of *phenotypic evolution*. On the other hand, the adaptiveness of the signalling system strategy can also be realised by psychological learning processes, as is seen with reinforcement learning models, which can be interpreted as a model of *cultural evolution*. In this case, a species learns to signal given some sort of positive (or negative) reinforcement. So, there are several different mechanisms by which the evolution of meaning may occur.

Regarding the second question—whether there are different interpretations of the models themselves—we get the same answer: one interpretation of the signalling game is that the population evolves according to some (deterministic) dynamic process, the other interpretation is that the population evolves according to some (stochastic) learning process. Even though these are different interpretations, it turns out that learning and evolution are closely related. Beggs (2005) and Hopkins and Posch (2005) showed that the mean-field dy-

[30] Orzack and Sober (1993) provide a critical analysis of the views in Levins (1966). See also the response to this in Levins (1993), and the general overview given in Weisberg (2006).

namic of Roth-Erev learning models (Roth and Erev, 1995; Erev and Roth, 1998) is a version of the replicator dynamic.[31]

The replicator dynamic itself has a dual interpretation—one biological and one cultural, and simple reinforcement learning, when mathematically formalised, produces a dynamic that is similar, and in some cases equivalent, to the replicator dynamic.[32] Thus, there is a certain generality to the dynamic framework upon which evolutionary models of signalling conventions are based which gives them an inherent flexibility in terms of both criteria.

4 The Sufficiency of Our Models

In the previous section, we have seen some considerations concerning the *adequacy* of evolutionary models of signalling. We turn now to the *sufficiency* question. With respect to the initial phenomenon to be explained, there are really two questions being asked here: (1) Are the models that we use sufficient to explain how and whether *signalling* arises spontaneously and is subsequently maintained in nature? (2) Are these models sufficient to explain the evolution of *language*. It is important to note that these are distinct questions—a signalling system may allow for meaningful communication, but this is not necessarily equivalent to *language* (e.g, human language). Nonetheless, signalling may be considered a *proto-language* which is a stepping stone to *language* proper, or it may be a stepping stone to *proto-language* which is itself a stepping stone toward language proper.

One thing to note is that signalling is ubiquitous in nature. Signalling can be observed in various species of monkey, including vervet monkeys (Cheney and Seyfarth, 1990), Diana monkeys (Zuberbühler, 2000), and Campbell's monkeys (Zuberbühler, 2001), as well as some species of Tamarins (Kirchhof and Hammerschmidt, 2006), lemurs (Macedonia, 1990), prairie dogs (Slobodchikoff et al., 1991), red squirrels (Greene and Maegher, 1998), etc. More impressively, perhaps, signalling can be observed in bees in the form of a complex dance (von Frisch, 1967); Black-capped chickadees have been observed to maintain a very strict syntactic structure in signalling (Hailman et al., 1985). Finally, we see the same simple signalling games can be used to model and explain behaviour as it arises in certain types of bacteria (Berleman et al., 2008).

Given that signalling is commonplace in nature, in order for our models to be sufficient, they must give results that show *why* and *how* this is so. The robustness of the results of the simplest case appear to give a sufficient explanation for the ubiquity of signalling in nature: signalling systems arise with near-certainty under a variety of (often representative) conditions. Additionally, since there were minimal assumptions (regarding natural salience,

[31] See also Schreiber (2001) for an analysis of the connection between the replicator dynamics and Pólya urns more generally.

[32] For more details on this, see Sandholm (2010).

correlation, intentionality, rationality, etc.), this is equally sufficient to explain signalling both in humans and monkeys, as well as in bees and bacteria.

This highlights the difference between language and signalling: signalling arises easily in nature and language apparently does not. One might argue that only humans have language, or even that only humans are *capable* of having language. However, even if we relaxed our criteria for considering something a language in the first place, it is still clear that language is not equivalent to signalling—if they were equivalent, then one would have to admit that quorum signalling in bacteria is indeed 'language'.

Evolutionary models seem, modulo some caveats raised in the discussion on adequacy, sufficient to explain the advent of simple signalling across various species. However, the progress made to date in this programme is still insufficient to explain the evolutionary jump from signalling to language.

One conceptual problem that falls within the purview of sufficiency is that despite the fact that evolutionary models of signalling games seem to show how meaning can arise in a population, it is not actually clear *what* that meaning is. There is an interpretive symmetry to be noted here: suppose a population already has a signalling system up and running wherein the strategy pairs in Figure 1a obtain, and there is near-perfect information transfer. The problem is that it is not clear whether the signal m_1 means the declarative 's_1 obtains' or the imperative 'do a_1'. This symmetry is pointed out by Harms (2004a,b) and is subsequently analysed by Huttegger (2007b) and Zollman (2011). Huttegger's model adds a 'deliberation' component to the simple signalling game wherein, if the receiver deliberates, then the signal is the declarative, whereas if the sender deliberates, then the signal is the imperative. However, Zollman points out that the same problem applies if we interpret the message as the declarative 'we are in s_1' or as the indicative 'deliberate, and then do a_1'. Zollman himself constructs a different model to explain how this symmetry gets broken—his model involves the addition of a third player. Zollman (2011) concludes by saying that

> I do not believe that the possibility of these two translations demonstrates an inadequacy of the Lewis signaling games as a model for linguistic behavior. Instead, this possibility is merely an example of the more radical problem of translation suggested by Quine['s *Word and Object*]. Merely because I can offer two different English sentences which both capture the meaning of the signals in the game is not sufficient to demonstrate that meaning is not present in these games[.] (168)

However, Harms' initial point is that there is not a translation of signals to English language expressions. Rather than trying to build a model that explains this lack of translation, one might just say (as Harms does) that a signal *qua* signalling is *primitive content*. Where 'primitive content' is similar to the *pushmi-pullyu* representations of teleosemantics (Millikan, 1995). So, the translation problem is evidence for the fact that signalling games do not constitute *languages*. That being said, if signalling games are sufficient for anything at this point, they are sufficient for explaining how signalling can arise from minimal assumptions, but not how *language* can arise from signalling.

Perhaps animals have not evolved past simple signalling because their biology is not naturally equipped for anything more complex—this may have to do with cognitive capacities, or it may have to do with physical capacities. One may argue that the reason humans evolved language where animals did not is simply because the right ingredients happened to be there: we may have evolved signalling like any other animal, but for some fortuitous reason we also had areas of the brain and anatomical tools like the larynx evolve to allow for the capacity of language. However, this will obviously be of no explanatory power here. In this way, sufficiency will require that future work focuses on closing the gap between signalling and language. Indeed, to some extent it has already begun to do so.

Barrett (2007, 2009) also introduces an extension of the simple Lewis signalling game wherein there are more players.[33] If there are two signals and 4 state/act pairs, perfect communication can still arise in the following way: there are two senders, each picks a signal from $\{0, 1\}$ and each individually sends her signal to the receiver. The receiver obtains each signal and picks an act. One interpretation of this model is that each sender has partial information, which combines to complete information for the receiver. Interpreted in this way, we can begin to model the evolution of simple logics—for example, if Sender 1 sends a signal meaning 's_1 or s_2' and Sender 2 sends a signal meaning '$\neg(s_1)$', then the receiver can infer s_2.[34]

Skyrms (2010a) took this 2-sender game and showed how it can be played with one player sending 2 signals in a particular order. This gives rise to the notion of a *syntactic* signalling game. Franke (2016) adds 'spill-over reinforcement' to this syntactic model and shows how something like compositionality might evolve. This last paper is particularly interesting because it does seem to be an effort to bridge the gap between signalling and language: Franke presupposes that a signalling system is *already up and running* and shows how something like compositionality can arise from that. These results, of course, would require the same critical treatment seen above. In this vein, Steinert-Threlkeld (2016) uses signalling games to explain why (instead of how) natural languages might have become compositional.

Additionally, Barrett (2016, 2017) has built a model of a meta-game which takes outcomes of a standard signalling game as input and evolves a metalinguistic notion of truth in parallel with the evolving language. Further, Barrett and Skyrms (2017) have built models in which complex signalling games evolve from several simple signalling games through *modular composition*.[35] Given the results of Spelke (2003), some notion of modularity seems to be of

[33] See also, Skyrms (2009, 2010a,b) for other cases of signalling networks.

[34] Note that a lot is being assumed here, so this may seem question-begging. However, Steinert-Threlkeld (2014) has shown that function words, such as 'not' may arise in signalling contexts.

[35] See also, Barrett et al. (2018); LaCroix (2019).

explanatory significance in further narrowing the gap between simple signals and language.

Pending further studies on the sufficiency and adequacy questions, signalling games at least have bridged the gap between the nonexistence and the existence of (primitive) meaning, couched in terms of communication as information transfer. Regarding the gap between signalling and language, the literature has seen advancements in explanations of proto-compositionality (Nowak and Krakauer, 1999; Barrett, 2006, 2007, 2009; Franke, 2016; Steinert-Threlkeld, 2016; Barrett et al., 2018), metalinguistic notions of truth and falsity (Barrett, 2017), the evolution of simple logical connectives, etc.[36] So, the simplest models of signalling are certainly sufficient to explain the existence of, e.g., quorum signalling in bacteria; further, though there is a large gap between quorum signalling systems and natural languages, progress is being made in bridging this gap.

5 Concluding Remarks and Future Research

We have seen now that signalling games perform reasonably well on all of the adequacy conditions: for representativeness, justifications can be made for nonrepresentative assumptions in the models; robustness is satisfied for all but the most stringent requirements for what counts as a success; and for flexibility, the simplest models perform well outright. Further, we saw that if the evolutionary models used for signalling games are to be sufficient, we need to differentiate what exactly they are supposed to be sufficient for. Note that there is *some* difference between signalling and language, but it is unclear precisely wherein that difference lies.[37] Nonetheless, simple signalling is certainly sufficient for explaining the simplest communication phenomena that we see in nature, though these models (and their underlying assumptions) will necessarily need to be extended and modified to account for how language might have arisen from simple communication.

The focus of this paper was the theoretical and empirical explanatory power of evolutionary models for the emergence of signalling. However, it should also be noted that there may be some practical application here to the advancement of artificial cognitive systems and computational intelligence. In order to recreate something like human intelligence, it is necessary to understand wherein that intelligence arises. Spelke (2003) has argued that 'what makes us smart' may be the way in which our different cognitive capacities combine through *modules*—complex cognitive structures made up of more simple 'core' knowledge systems. It is argued then that the difference in cognitive capacities between human and non-human animals obtains in the combinatorial capacities that human cognitive systems have that allow them to conjoin various

[36] See also the iterated learning model (Kirby and Hurford, 2002; Smith et al., 2003), which is connected to the signalling game framework by Spike et al. (2013).

[37] Most researcher hold that key distinction between language and animal communication systems is that the former utilises compositional syntax. See Progovac (2019) for an overview.

representations of the world in order to *create* new cognitive knowledge. Furthermore, it is commonly believed that there is a close relationship between language and thought. As such, a clear understanding of how languages arise and are maintained in a population can give us some further insight into human cognitive capacities. The possibility to model and understand these sort of linguistic systems may allow for more sophisticated replication of human cognitive capacities in artificial intelligence programs, for example. So, even the simplest of these models may additionally have some practical significance.

References

Ahern C, Clark R (2014) Diachronic processes in language as signaling under conflicting interests. In: Cartmill E (ed) The Evolution of Language: Proceedings of the Tenth International Conference on the Evolution of Language, Vienna, pp 25–32

Alexander JM (2007) The Structural Evolution of Morality. Cambridge University Press, Cambridge

Alexander JM, Skyrms B, Zabell SL (2012) Inventing New Signals. Dynamic Games and Applications 2:129–145

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

Barrett JA (2006) Numerical Simulations of the Lewis Signaling Game: Learning Strategies, Pooling Equilibria, and Evolution of Grammar. Technical Report, Institute for Mathematical Behavioral Science

Barrett JA (2007) Dynamic Partitioning and the Conventionality of Kinds. Philosophy of Science 74:527–546

Barrett JA (2009) The Evolution of Coding in Signaling Games. Theory and Decision 67:223–237

Barrett JA (2016) On the Evolution of Truth. Erkenntnis 81:1323–1332

Barrett JA (2017) Truth and Probability in Evolutionary Games. Journal of Experimental and Theoretical Artificial Intelligence 29(1):219–225

Barrett JA, Skyrms B (2017) Self-Assembling Games. The British Journal for the Philosophy of Science 68(2):329–353

Barrett JA, Zollman K (2009) The Role of Forgetting in the Evolution and Learning of Language. Journal of Experimental and Theoretical Artificial Intelligence 21(4):293–309

Barrett JA, Skyrms B, Cochran C (2018) Hierarchical Models for the Evolution of Compositional Language, URL http://philsci-archive.pitt.edu/14725/, unpublished manuscript. May, 2018. PDF File

Batterman RW (2009) Idealization and Modeling. Synthese 169(3):427–446

Beggs AW (2005) On the Convergence of Reinforcement Learning. Journal of Economic Theory 122:1–36

Bereby-Meyer Y, Erev I (1998) On Learning to Become a Successful Loser: A Comparison of Alternative Abstractions of Learning Processes in the Loss Domain. Journal of Mathematical Psychology 42(2):266–286

Berleman JE, Scott J, Chumley T, Kirby JR (2008) Predataxis Behaviour in *Myxococcus Xanthus*. In: Proceedings of the National Academy of Sciences USA, pp 17127–17132

Bratman ME (1999) Faces of Intention. Cambridge University Press, Cambridge

Bruner J, O'Connor C, Rubin H, Huttegger SM (2018) David Lewis in the Lab: Experimental Results on the Emergence of Meaning. Synthese 195(2):603–621

Brusse C, Bruner J (2017) Responsiveness and Robustness in the David Lewis Signaling Game. Philosophy of Science 84(5):1068–1079

Bush RR, Mosteller F (1955) Stochastic Models for Learning. John Wiley & Sons, Oxford

Cheney D, Seyfarth R (1990) How monkeys See the World: Inside the Mind of Another Species. University of Chicago Press, Chicago

Clark HH (1996) Using Language. Cambridge University Press, Cambridge

Cohen PR, Levesque HJ (1991) Teamwork. Nôus 25:487–512

Crawford VP, Sobel J (1982) Strategic Information Transmission. Econometrica 50(6):1431–1451

D'Arms J (1996) Sex, Fairness, and the Theory of Games. Journal of Philosophy 93(12):615–627

D'Arms J (2000) When Evolutionary Game Theory Explains Morality, What Does It Explain? Journal of Consciousness Studies 7(1–2):296–299

D'Arms J, Batterman R, Gorny K (1998) Game Theoretic Explanations and the Evolution of Justice. Philosophy of Science 65(1):76–102

Donaldson MC, Lachmann M, Bergstrom CT (2007) The Evolution of Functionally Referential Meaning in a Structured World. Journal of Theoretical Biology 246:225–233

Erev I, Roth AE (1998) Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. American Economic Review 88(4):848–881

Fitch WT (2010) The Evolution of Language. Cambridge University Press, Cambridge

Fitch WT (2017) Empirical Approaches to the Study of Language Evolution. Psychonomic Bulletin & Review 24(1):3–33

Franke M (2016) The Evolution of Compositionality in Signaling Games. Journal of Logic, Language and Information 25(3):355–377

Franke M, Wagner EO (2014) Game Theory and the Evolution of Meaning. Language and Linguistics Compass 8/9:359–372

Franke M, de Jager T, van Rooij R (2012) Relevance in Cooperation and Conflict. Journal of Logic and Computation 22(1):23–54

von Frisch K (1967) The Dance Language and Orientation of the Bees. Harvard University Press, Cambridge, MA

Gilbert M (1989) On Social Facts. Princeton University Press, Princeton

Gintis H (2000) Classical Versus Evolutionary Game Theory. Journal of Consciousness Studies 7(1–2):300–304

Godfrey-Smith P, Martínez M (2013) Communication and Common Interest. PLoS Computational Biology 9(11):1–6

Greene E, Maegher T (1998) Red Squirrels, *Tamiasciurus hudsonicus*, Produce Predator-Class Specific Alarm Calls. Animal Behaviour 55:511–518

Hailman J, Ficken M, Ficken R (1985) The 'Chick-a-dee' calls of *Parus atricapillus*. Semiotica 56:191–224

Harms WF (2000) The Evolution of Cooperation in Hostile Environments. Journal of Consciousness Studies 7(1–2):308–313

Harms WF (2004a) Information and Meaning in Evolutionary Processes. Cambridge University Press, Cambridge

Harms WF (2004b) Primitive content, translation, and the emergence of meaning in animal communication. In: Oller DK, Griebel U (eds) Evolution of Communication Systems: A Comparative Approach, MIT Press, Cambridge, MA, pp 31–48

Herrnstein RJ (1970) On the Law of Effect. Journal of Experimental Analysis of Behavior 13:243–266

Hofbauer J, Huttegger S (2008) The Feasibility of Communication in Binary Signaling Games. Journal of Theoretical Biology 254:843–849

Hofbauer J, Sigmund K (1998) Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge

Hofbauer J, Sigmund K (2003) Evolutionary Game Dynamics. Bulletin of the American Mathematical Society 40(4):479–519

Hopkins E, Posch M (2005) Attainability of Boundary Points Under Reinforcement Learning. Games and Economic Behavior 53:110–125

Humphreys P, Imbert C (2012) Models, Simulations, and Representations, Routledge Studies in the Philosophy of Science, vol 9. Routledge, New York

Huttegger SM (2007a) Evolution and the Explanation of Meaning. Philosophy of Science 74:1–27

Huttegger SM (2007b) Evolutionary Explanations of Indicatives and Imperatives. Erkenntnis 66:409–436

Huttegger SM, Skyrms B, Smead R, Zollman KJS (2010) Evolutionary Dynamics of Lewis Signaling Games: Signaling Systems vs. Partial Pooling. Synthese 172(1):177–191

Jackendoff RS (1999) Possible Stages in the Evolution of the Language Capacity. Trends in Cognitive Sciences 3:272–279

Jackendoff RS (2002) Foundations of Language: Brain, Meaning, Grammar, Evolution. Oxford University Press, Oxford

Kirby S, Hurford JR (2002) The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model. In: Cangelosi A, Parisi D (eds) Simulating the Evolution of Language, Springer, London, pp 121–147

Kirchhof J, Hammerschmidt K (2006) Functionally Referential Alarm Calls in Tamarins (Saguinus fuscicollis and Saguinus mystax)—Evidence From Playback Experiments. Ethology 112:346–354

Kitcher P (1999) Games Social Animals Play: Commentary on Brian Skyrms' *Evolution of the Social Contract*. Philosophy and Phenomenological Research 59(1):221–228

LaCroix T (2018) On Salience and Signaling in Sender-Receiver Games: Partial Pooling, Learning, and Focal Points. Synthese Forthcoming

LaCroix T (2019) Using logic to evolve more logic: Composing logical operators via self-assembly, unpublished manuscript. May, 2019. PDF File

Levins R (1966) The Strategy of Model Building in Population Biology. In: Sober E (ed) Conceptual Issues in Evolutionary Biology, MIT Press, Cambridge, MA, pp 18–27

Levins R (1993) A Response to Orzack and Sober: Formal Analysis and the Fluidity of Science. Quarterly Review of Biology 68(4):547–555

Lewis D (2002/1969) Convention. Blackwell, Oxford

Lieberman P (2000) Human Language and Our Reptilian Brain: The Subcortical Bases of Speech, Syntax and Thought. Harvard University Press, Cambridge, MA

Macedonia JM (1990) What is Communicated in the Antipredator Calls of Lemurs: Evidence from Antipredator Call Playbacks to Ring-Tailed and Ruffed Lemurs. Ethology 86(3):177–190

Martínez M, Godfrey-Smith P (2016) Common Interest and Signaling Games: A Dynamic Analysis. Philosophy of Science 83(3):371–392

Maynard Smith J (1991) Honest Signalling: The Philip Sidney Game. Animal Behaviour 42:1034–1035

Mayr E (1961) Cause and Effect in Biology. Science 134(3489):1501–1506

Millikan RG (1995) Pushmi-Pullyu Representations. Philosophical Perspectives 9:185–200

Mühlenbernd R (2011) Learning with neighbours: Emergence of Convention in a Society of Learning Agents. Synthese 183:87–109

Mühlenbernd R, Franke M (2014) Meaning, Evolution, and the Structure of Society, cEUR Workshop Proceedings

Mühlenbernd R, Nick JD (2014) Language change and the force of innovation. In: Katrenko S, Rendsvig RK (eds) Pristine Perspectives on Logic, Language, and Computation, no. 8607 in Lecture Notes in Computer Science, Springer, Heidelberg/New York, pp 194–213

Nash J (1951) Non-Cooperative Games. The Annals of Mathematics 54(2):286–295

Neumann JV, Morgenstern O (2007/1944) Theory of Games and Economic Behavior. Princeton university press

Nowak MA (2006) Evolutionary Dynamics: Exploring the Equations of Life. The Belknap Press of Harvard University Press, Cambridge, MA

Nowak MA, Krakauer DC (1999) The Evolution of Language. Proceedings of the National Academy of Sciences 96:8028–8033

Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of Cooperation and Evolutionary Stability in Finite Populations. Nature 428:646–650

Orzack SH, Sober E (1993) A Critical Assessment of Levins's 'The Strategy of Model Building in Population Biology' (1966). Biology and Philosophy

68(4):533–546
Pawlowitsch C (2007) Finite Populations Choose an Optimal Language. Journal of Theoretical Biology 249:606–616
Pawlowitsch C (2008) Why Evolution Does Not Always Lead to an Optimal Signaling System. Games and Economic Behavior 63(1):203–226
Pinker S, Bloom P (1990) Natural Language and Natural Selection. Behavioral and Brain Sciences 13:707–726
Pinker S, Jackendoff R (2005) The Faculty of Language: What's Special About It? Cognition 95:201–236
Progovac L (2015) Evolutionary Syntax. Oxford University Press, Oxford
Progovac L (2019) A Critical Introduction to Language Evolution: Current Controversies and Future Prospects. Springer, Berlin
Quine WVO (1967) Truth by Convention. In: Philosophical Essays for Alfred North Whitehead, Russell & Russell, New York, pp 90–124
Roth AE, 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
Russell B (1922) Analysis of Mind. The MacMillan Company, New York
Sandholm WH (2010) Population Games and Evolutionary Dynamics. The MIT Press, Cambridge, MA
Santana C (2014) Ambiguity in Cooperative Signaling. Philosophy of Science 81(3):398–422
Schelling TC (1980/1960) The Strategy of Conflict. Harvard University Press, Cambridge, MA
Schreiber SJ (2001) Urn Models, Replicator Processes, and Random Genetic Drift. SIAM Journal on Applied Mathematics 61(6):2148–2167
Searle J (1995) The Construction of Social Reality. Free Press, New York
Skyrms B (1990) The Dynamics of Rational Deliberation. Harvard University Press, Cambridge
Skyrms B (1994) Darwin Meets the Logic of Decision: Correlation in Evolutionary Game Theory. Philosophy of Science 61(4):503–528
Skyrms B (2000) Stability and Explanatory Significance of Some Simple Evolutionary Models. Philosophy of Science 67(1):94–113
Skyrms B (2004) The Stag Hunt and the Evolution of Social Structure. Cambridge University Press, Cambridge
Skyrms B (2009) Evolution of Signaling Systems with Multiple Senders and Receivers. Philosophical Transactions of the Royal Society B 364:771–779
Skyrms B (2010a) Signals: Evolution, Learning, & Information. Oxford University Press, Oxford
Skyrms B (2010b) The Flow of Information in Signaling Games. Philosophical Studies 147(1):155–165
Skyrms B (2014/1996) Evolution of the Social Contract. Cambridge University Press, Cambridge
Slobodchikoff CN, Kiriazis J, Fischer C, Creef E (1991) Semantic information distinguishing individual predators in the alarm calls of gunnison's prairie dogs. Animal Behaviour 42:713–719

Smith K, Kirby S, Brigthon H (2003) Iterated learning: A framework for the emergence of language. Artificial Life 9(4):371–386

Spelke E (2003) What Makes Us Smart? Core Knowledge and Natural Language. In: Gentner D, Goldin-Meadow S (eds) Language in Mind: Advances in the Investigation of Language and Thought, MIT Press, Cambridge, MA, pp 277–311

Spence M (1973) Job market signaling. The Quarterly Journal of Economics 87(3):355–374

Spike M, Stadler K, Kirby S, Smith K (2013) Learning, feedback and information in self-organizing communication systems. Proceedings of Cogsci 2013 87(3):3442–3447

Steinert-Threlkeld S (2014) Learning to use function words in signaling games. In: Lorini E, Perrussel L (eds) Proceedings of *Information Dynamics in Artificial Societies*, (IDAS-14)

Steinert-Threlkeld S (2016) Compositional signaling in a complex world. Journal of Logic, Language, and Information 25(3–4):379–397

Sterelny K (2012) A glass half-full: Brian skyrms's *signals*. Economics and Philosophy 28(1):73–86

Sugden R (2000) Credible worlds: The status of theoretical models in economics. Journal of Economic Methodology 1:1–31

Taylor P, Jonker L (1978) Evolutionarily Stable Strategies and Game Dynamics. Mathematical Biosciences 40:145–156

Thorndike EL (1905) The Elements of Psychology. The Mason Press, Syracuse

Thorndike EL (1911) Animal Intelligence: Experimental Studies. The Macmillan Company, New York

Thorndike EL (1927) The Law of Effect. American Journal of Psychology 39:212–222

Tomasello M (1999) The Cultural Origins of Human Cognition. Harvard University Press, Cambridge, MA

Vanderschraaf P (1995) Convention as Correlated Equilibrium. Erkenntnis 42:65–87

Wagner EO (2009) Communication and Structured Correlation. Erkenntnis 71(3):377–393

Weisberg M (2006) Robustness Analysis. Philosophy of Science 73(5):730–742

Wimsatt WC (2007) Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality. Harvard University Press, Cambridge, MA

Zeeman EC (1980) Population Dynamics From Game Theory. In: Global Theory of Dynamical Systems, Lecture Notes in Mathematics, vol 819, Springer

Zollman KJS (2005) Talking to Neighbors: The Evolution of Regional Meaning. Philosophy of Science 72(1):69–85

Zollman KJS (2011) Separating Directives and Assertions Using Simple Signaling Games. The Journal of Philosophy 108(3):158–169

Zuberbühler K (2000) Referential Labeling in Diana Monkeys. Animal Behaviour 59:917–927

Zuberbühler K (2001) Predator-Specific Alarm Calls in Campbell's Monkeys, *Cercopithecus Campbelli*. Behavioural Ecology and Sociobiology 50:414–422