BRIAN SKYRMS

SALIENCE AND SYMMETRY-BREAKING IN THE EVOLUTION OF CONVENTION[1]

Since monkeys certainly understand much that is said to them by man, and when wild, utter signal-cries of danger to their fellows; and since fowls give distinct warnings for danger on the ground, or in the sky from hawks (both, as well, a third cry, intelligible to dogs), may not some unusually wise ape-like animal have imitated the growl of a beast of prey, and thus told his fellow-monkeys the nature of the expected danger? This would have been the first step in the formation of a language.

Charles Darwin *The Descent of Man*

The first step in the formation of language is the first step in the formation of a convention. Predator-specific alarm calls can serve a useful purpose only if the animal giving a call and the animal receiving the call associate types of predators with types of calls in the same way. There are always many alternative patterns of association that would serve to transmit the information. How is a particular pattern initially selected? Darwin's answer is that selection is effected by *natural salience*. The lookout imitates the "growl of the beast of prey" and his fellows who hear this imitation are naturally led to the correct conclusion.

Contemporary studies confirm the existence of predator-specific alarm calls in various species, including vervet monkeys, ring-tailed lemurs, superb starlings and domestic chickens. However, the current literature on alarm calls fails to confirm the role of the kind of natural salience that Darwin had in mind. Alarm calls do not seem to be imitations of the calls of the associated predator. Other considerations seem to be in play:

[1] The subjects of this paper are treated at greater length in Chapters 4 and 5 of my 1996 book. See Skyrms, *Evolution of the Social Contract* (New York: Cambridge University Press, 1996).

Law and Philosophy **17:** 411–418, 1998.

> Vervets on the Cameroon savanna are sometimes attacked by feral dogs. When they see a dog they respond much as Amboseli vervets respond to a leopard; they give loud alarm calls and run into trees. Elsewhere in the Cameroon, however, vervets live in forests where they are hunted by armed humans who track them down with the aid of dogs. In these circumstances, where loud alarm calls and conspicuous flight into trees would only increase the monkey's likelihood of being shot, the vervets alarm calls to dogs are short, quiet, and cause others to flee silently into the dense bush where humans cannot follow.[2]

There is evidence that the hawk alarm of many small birds has characteristics which make it hard for the hawk to locate.[3]

This is not to say that natural salience did not play a recognizable part in the evolution of other animal signals, involving threat or mating. It did. And we cannot completely rule out the possibility that species-specific alarm calls arose as Darwin conjectured, but were changed in subsequent evolution to the extent that their origins are unrecognizable. But lacking any evidence for Darwin's conjecture, we are led to ask whether a system of alarm signals could have arisen in the absence of natural salience.

The question can be put cleanly within the framework of sender-receiver signaling games of the kind introduced in David Lewis' study of convention.[4] The sender gets private information about the world which he wishes to communicate to the receiver. At his disposal, he has some signals which have no preexisting meaning. He sends a signal to the receiver, who then chooses among alternative acts. Sender and receiver are supposed to have common interest. If the sender succeeds in communicating the information to the receiver, the receiver chooses the act that is best for both of them.

---

[2] D. L. Cheney and R. M. Seyfarth, *How Monkeys See the World: Inside the Mind of Another Species* (Chicago: University of Chicago Press, 1990), p. 169, citing M. Kavanaugh, "Invasion of the Forest by an African Savanah Monkey: Behavioral Adaptations," *Behavior* 73 (1980).

[3] See, e.g., P. Marler, "Characteristics of some animal calls," *Nature* 176 (1955), pp. 6–7 and C. H. Brown, "Ventrologuial and locatable vocalizations in birds," *Zeitschrift für Tierpsychologie* 59 (1982), pp. 338–350.

[4] Lewis' account of signaling games is generalized by Crawford and Sobel. See D. K. Lewis, *Convention: A Philosophical Study* (Cambridge, Mass.: Harvard University Press, 1969); and V. P. Crawford and J. Sobel, "Strategic Information Transmission," *Econometrica* 50 (1982), pp. 1431–1451.

Consider the alarm calls of the Amboseli vervet monkeys studied by Cheney and Seyfarth.[5] The major predators that they face are leopards, eagles and snakes. Each has a different mode of attack, and for each there is a different optimal evasive strategy. Abstracting from the situation we are led to consider a game where the sender observes one of three possible states of the world: S1, S2, S3 and sends one of three possible messages M1, M2, M3. The receiver then takes one of three possible acts, A1, A2, A3. In each state of the world exactly one act is optimal for both players. We number the acts and states for convenience so that if the number of the act that is taken matches the number of the state of the world, both players get a payoff of 1, otherwise both get a payoff of 0.

A sender's strategy is a rule telling the sender what message to send in what state of the world; a receiver's strategy is a rule mapping *message received* onto *action to take*. If the sender's strategy and the receiver's strategy fit together so that the receiver always makes the choice that is best in the state observed by the sender, Lewis says that these strategies taken together constitute a *signaling system*. The vervets have evolved a signaling system, but it is not the only possible one. Any permutation of signals gives us another possible signaling system. At least in our abstract model, one signaling system is as good as another. And there are many game-theoretic equilibria (Nash equilibria) in these games that are not signaling systems at all. Could the vervets have coordinated on a particular signaling system in the absence of natural salience? How could their signaling convention have evolved?

For the answer, we must look at the underlying evolutionary dynamics. Let us let all kinds of combination of sender and receiver strategies arise in the population. Start with some population proportions picked at random and examine the results of differential reproduction on that population over time. Then choose new population proportions at random and repeat the process. I ran such a computer simulation, with the result that signaling systems *always* evolved. This happened in a model where the possibility that one signaling system equilibrium was especially salient was deliberately excluded. In fact, the signaling system which evolved was not always the same. Each possible signaling system evolved on some

[5] Cheney and Seyfarth, *How Monkeys See the World*, Chapter 4.

of the trials. The equilibria which are not signaling systems never evolved. The reason for this is that they are dynamically unstable. Only signaling systems are attractors in the evolutionary dynamics.[6]

I want to emphasize here that I am not saying that natural salience never plays a part in evolutionary equilibrium selection. I am only making the point that in situations where the dynamical structure induced by the problem is favorable, natural salience may be unnecessary for the emergence of a convention.

The point is not confined to strictly evolutionary settings. The structure of Lewis' signaling games suggests that other adaptive dynamics – for instance dynamics of group learning – ought to deliver comparable results. This hypothesis has been put to the test by experimental game theorists. Blume, Kim and Sprinkle saw whether undergraduates at the University of Iowa would spontaneously learn to play some signaling system in a sender-receiver game of the kind discussed by Lewis.[7]

They take extraordinary precautions to exclude natural salience from the experimental setting. Sender and receiver communicate to each other over a computer network. The messages available to the sender are the asterisk and the pound sign, {*, #}. These are identified to the players as possible messages on their computer screens. The order in which they appear on a given player's screen is chosen at random to control for the possibility that order of presentation might function as the operative salience cue. Then players repeatedly play a Lewis signaling game. Players are kept informed of the history of play of the group. (The group consists of 6 senders and 6 receivers. After each play of the signaling game, the players are updated on what happened.) Under these conditions the players rapidly (in 15 or 20 periods) learn to coordinate on one signaling system or another.

In these games of common interest, evolutionary dynamics, learning dynamics, and almost any reasonable sort of adaptive dynamics leads to successful coordination on a signaling system equilibrium. In the absence of natural salience, which signaling

---

[6] I use the replicator dynamics for these simulations. More details may be found in Skyrms, *supra* note 1.

[7] A. Blume, Y-G. Kim, and G. B. Sprinkle, "Evolution of the Meaning of Messages in Sender-Receiver Games" (1996) working paper, University of Iowa.

system emerges depends on the vicissitudes of initial conditions and chance aspects of the process. But some signaling system does evolve because signaling systems are powerful attractors in the dynamics, and other Nash equilibria of the game are dynamically unstable.

Let us turn our attention now to a game that is not a game of common interest, but rather one that has competitive and cooperative aspects. This is the game known to economists and James Dean fans as "Chicken" and to evolutionary biologists as "Hawk-Dove".[8] Two animals, perhaps human, contest a resource. Each implements one of two strategies: Hawk or Dove – whose names are self-explanatory. Hawks beat Doves, but when two Hawks meet they both suffer whereas when two doves meet they coexist. In terms of reward, it is best to be a Hawk playing against a Dove, next best to be a Dove Playing against a Dove, next best to be a Dove playing against a Hawk, and worst to be a Hawk playing against another Hawk.[9]

Even though Hawks always beat Doves, they do not take over the population. Assuming random encounters, in a population with mostly Hawks, it is better to be a Dove. In a population with mostly Doves, it is better to be a Hawk. Evolutionary dynamics drives the population to a polymorphic state with some Hawks and some Doves.[10]

This does not seem to be a state propitious for the evolution of convention. But notice that this polymorphism may be inefficient for all concerned.[11] The feature of the evolutionary context that is responsible for this sorry state of affairs is the symmetrical position of the two strategies. Players do not persist over evolutionary time. The objects of evolution are strategies. The pure equilibria of rational choice game theory: I play Hawk and you play Dove; you play Hawk and I play Dove are not options because you and I are not

[8] J. Maynard Smith and G. R. Price, "The Logic of Animal Conflict," *Nature* 146 (1973), pp. 15–18.

[9] For a numerical example, suppose payoff of Hawk against Dove is 50, of Dove against Dove is 15, of Dove against Hawk is 0 and of Hawk against Hawk is -25.

[10] In the example of the previous footnote, the equilibrium state has 5/12 Doves and 7/12 Hawks.

[11] Average payoff here in our numerical example is just 6 1/4.

significant in evolutionary time – even when we are talking about cultural evolution. The *symmetry* of the players' positions forces the population to a highly undesirable state.

There may, however, be ways in which symmetry can be broken. Suppose, however, that there is some random feature of the environment which is apparent to the players and on which they could condition their actions. Schematically, let us say that nature spins a pointer and it points to one of the players, and the players notice this. Then, additional strategies become possible. We have: *Hawk if the pointer points to me; Dove otherwise*; *Dove if the pointer points to me, Hawk otherwise*, as well as: *Hawk regardless* and Dove *regardless*. The pointer opens up the possibility that the evolutionary symmetry may be broken. Evolutionary dynamics assures that it will be broken. For now the basins of attraction of the two conditional strategies span almost all the space of states, and the polymorphic equilibrium state of Hawks and Doves becomes dynamically unstable.

The dynamical analysis predicts that either the population evolves either to the state where everyone plays *Hawk if the pointer points to me, Dove otherwise* or to the state where everyone plays *Dove if the pointer points to me, Hawk otherwise.* (The office of the pointer could be discharged by various sorts of events – for example, by who got there first.) Each of these states is a correlated convention, in the sense Vanderschraaf discusses.[12]

In a certain sense, salience is built into the foregoing model, but it is salience of the random event rather that salience of a particular *equilibrium*. There are, after all, two equally salient equilibria which employ strategies that condition on the random event. It is not the case that one or the other is *focal* in Schelling's sense.[13]

The random event is something that agents must be able to notice, but given that they can notice many things, why should they be drawn to condition their strategies on the random event in question? Of course, if one agent were to believe that other agents conditioned

[12] P. Vanderschraaf, "Convention as Coordinated Equilibrium," *Erkenntnis* 42 (1995), pp. 65–87. See also the discussion of one-sided signaling systems in Lewis, *Convention*, pp. 128–130.

[13] Schelling, *The Strategy of Conflict* (Cambridge, Mass.: Harvard University Press, 1969), pp. 68–70, 111–113.

their strategies on this random event, it would a salient event for her, in the special sense of salience now under discussion. But this seems to beg the question, for she may believe that other agents, like herself, have no special reason to attribute strategic significance to this event.

For agents capable of learning, we can give a dynamical resolution of this paradox. Even though agents are initially skeptical about the strategic relevance of the random event, as long as they allow for the possibility learning dynamics may make it relevant. Suppose that there are a number of random events that the agents may notice. By chance, in a finite sample there may arise apparent small correlations between the outcomes some random event, and the player's actions. Learning dynamics amplifies the correlation. A spurious correlation becomes a real correlation. When players believe that the random event is relevant, they condition their actions on it, and it becomes relevant to their interaction. Which random event or events become salient in this way is itself a matter of chance, depending on the timing of apparent correlations and learning. The essential point here is that when the circumstances of interaction are favorable, the dynamics of learning can function both as a salience amplifier and as a spontaneous salience generator.[14]

The external random event need not be objectively random in any stringent sense. All that is required is that it appear random to the players. For example, Maynard Smith and Parker consider strategies which break the symmetry of the Hawk-Dove game depending on who is the "owner" and who is the "intruder". And contemporary drivers (usually) condition their strategies at intersections on the color of the traffic light.

Thus the dynamical account of the selection of a convention can be extended from signaling systems, which are games of common interest, to games like Hawk-Dove, which are not. The players can use an external random event to break the symmetry of an inefficient mixed equilibrium, and move to an efficient correlated equilibrium. One may, with Lewis,[15] think of the external random

[14] For a more detailed discussion of this point, see P. Vanderschraaf and B. Skyrms, "Deliberation and Coordinated Equilibria," *Philosophical Topics* 21 (1993), pp. 191–227.

[15] See Lewis' discussion of the stoplight, *Convention*, p. 129.

event as functioning as a “signal from nature”. In some cases, a society will find it expedient to construct a device to reliably produce the “signal from nature”, as in the case of traffic lights. Here, salience of the signal and of the selected signaling system equilibrium, derives explicitly from the law. On the other hand, in cases where there is no preexisting salience, salience can be spontaneously produced by the dynamics of learning.

Humanities and Social Science 228-77
California Institute of Technology
Pasadena, CA 91125
U.S.A.