EVOLUTION, NORMS, AND THE SOCIAL CONTRACT

Brian Skyrms*

INTRODUCTION

I think of the social contract not as some monolithic unitary pact, but as an assemblage of norms. Norms are conventions that are backed by sanctions. Sometimes the sanctions are codified in the law and enforced by government. Sometimes norms are not explicit, but rather implicit in practice, and sanctions take the form of some type of social pressure.

Conventions, and thus norms, evolve, as do the sanctions that sometimes arise to back them. It is therefore useful to explore issues of the social contract using tools of evolutionary game theory, where "evolution" includes cultural evolution. This was all seen clearly, before evolutionary game theory even existed, by the great philosopher David Hume:[1]

> Two men, who pull the oars of a boat, do it by an agreement or convention, tho' they have never given promises to each other. Nor is the rule concerning the stability of possession the less deriv'd from human conventions, that it arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it. . . . In like manner are languages gradually establish'd by human conventions without any promise. In like manner do gold and silver become the common measures of exchange, and are esteem'd sufficient payment for what is of a hundred times their value.[2]

---

* Distinguished Professor of Logic and Philosophy of Science at the University of California, Irvine and Professor of Philosophy at Stanford University. I would like to thank Robin Bradley Kar and the Society for Evolutionary Analysis in Law for the kind invitation to contribute a keynote to the 16th SEAL conference.

1. *See, e.g.*, Peter Vanderschraaf, *The Informal Game Theory in Hume's Account of Convention*, 14 ECON. & PHIL. 215, 215–17 (1998). For modern accounts of convention using rational choice game theory, see DAVID LEWIS, CONVENTION: A PHILOSPHICAL STUDY (1969); Peter Vanderschraaf, *Convention as Correlated Equilibrium*, 42 ERKENNTNIS 65, 65–67 (1995) [hereinafter Vanderschraff, *Correlated Equilibrium*].

2. DAVID HUME, A TREATISE OF HUMAN NATURE 490 (L.A. Selby-Bigge ed., rev. 3d ed. 1976).

Evolutionary game theory arose in modern form, fusing game theory and evolutionary dynamics, in the twentieth century.[3] The founding fathers were William Hamilton, George Price, and John Maynard Smith.[4] It is now an established, growing subject with results that fill textbooks.[5]

I have been working in collaboration with friends on evolution of conventions, norms, and the social contract for some time, and what I can present here are just some of the themes developed in a series of books and papers.[6] Closely related work has been pursued by economists such as Ken Binmore,[7] Robert Sugden,[8] and Peyton Young.[9] There is also related work in philosophy by Cristina Bicchieri,[10] Peter Vanderschraaf,[11] Gregory Kavka,[12] Jean Hampton,[13] J. McKenzie Alexander,[14] and others.

I. THE NATURE OF THE SOCIAL CONTRACT

Traditional, and most of contemporary, social contract theory has a quite different character than an evolutionary theory of the social contract. Traditional theory asks what kind of contract would have been reached by rational, reasonable agents if they were in the position of setting up an ideal

3. *See* W.G.S. Hines, *Evolutionary Stable Strategies: A Review of Basic Theory*, 31 THEORETICAL POPULATION BIOLOGY 195, 195 (1987).

4. *See* Karl Sigmund, *William D. Hamilton's Work in Evolutionary Game Theory*, 59 THEORETICAL POPULATION BIOLOGY 3, 3 (2001).

5. *See, e.g.*, JOSEF HOFBAUER & KARL SIGMUND, EVOLUTIONARY GAMES AND POPULATION DYNAMICS (1998); WILLIAM SANDHOLM, POPULATION GAMES AND EVOLUTIONARY DYNAMICS (2010); JORGEN W. WEIBULL, EVOLUTIONARY GAME THEORY (1995).

6. BRIAN SKYRMS, EVOLUTION OF THE SOCIAL CONTRACT (2d ed. 2014); BRIAN SKYRMS, SIGNALS: EVOLUTION, LEARNING AND INFORMATION (2010) [hereinafter SKYRMS, SIGNALS]; BRIAN SKYRMS, SOCIAL DYNAMICS (2014); BRIAN SKYRMS, THE STAG HUNT AND THE EVOLUTION OF SOCIAL STRUCTURE (2004) [hereinafter SKYRMS, STAG HUNT].

7. *See* KEN BINMORE, GAME THEORY AND THE SOCIAL CONTRACT, VOLUME 1: PLAYING FAIR (1994) [hereinafter BINMORE, PLAYING FAIR]; KEN BINMORE, GAME THEORY AND THE SOCIAL CONTRACT, VOLUME 2: JUST PLAYING (1998) [hereinafter BINMORE, JUST PLAYING].

8. *See* ROBERT SUGDEN, THE ECONOMICS OF RIGHTS, COOPERATION AND WELFARE (2005).

9. H. PEYTON YOUNG, INDIVIDUAL STRATEGY AND SOCIAL STRUCTURE (1998); H. Peyton Young, *The Economics of Convention*, 10 J. ECON. PERSP. 105 (1996); H. Peyton Young, *The Evolution of Conventions*, 61 ECONOMETRICA 57 (1993).

10. CRISTINA BICCHIERI, THE GRAMMAR OF SOCIETY: THE NATURE AND DYNAMICS OF SOCIAL NORMS (2006).

11. PETER VANDERSCHRAAF, LEARNING AND COORDINATION: INDUCTIVE DELIBERATION, EQUILIBRIUM, AND CONVENTION (2001).

12. GREGORY KAVKA, HOBBESIAN MORAL AND POLITICAL THEORY (1986).

13. JEAN HAMPTON, HOBBES AND THE SOCIAL CONTRACT TRADITION (1986).

14. J. MCKENZIE ALEXANDER, THE STRUCTURAL EVOLUTION OF MORALITY (2007).

contract.[15] It is more or less assumed that all such agents would come to the same conclusion, and that gives the one, ideal social contract.[16] The contract settles all sorts of questions. It is thought to carry some justificatory force.

Skeptics, since ancient times, have pointed out that different cultures have arrived at different social contracts. This is dismissed as saying that some, or perhaps all, are not rational or not reasonable. A contemporary skeptic might point out that one leading theorist at a leading institution of higher learning may arrive at one contract, while another leading theorist with an office down the hall might arrive at another, while each maintains that any rational reasonable person would agree with his view.

In contrast, the evolutionary view expects multiple equilibria.[17] The vicissitudes of history can lead to different social contracts. This calls for scrutiny of the dynamics of evolution.[18] For cultural evolution, there is a whole range of dynamics that can be relevant. Dynamics may incorporate various kinds and degrees of bounded rationality. Dynamics play a major role in selecting equilibria. In certain circumstances dynamics may never reach equilibrium.[19] If we may never be at equilibrium, and if evolutionary dynamics may not respect perfect rationality, we might well question the emphasis on equilibrium and perfect rationality in traditional social contract theory. If we add in the insight that the social contract does not simply deal with one kind of social interaction, one game, but with many kinds—that it is a conglomerate of norms—then the traditional theory seems to aim in principle at a vast oversimplification.

Does a naturalistic evolutionary contract theory deliver an account of what we ought to do, or agree to, or approve as rational, reasonable beings? No it does not. But, all pretense aside, traditional theories do not either.

II. SOME GAMES

Game theoretic discussions of the social contract often focus on one specific paradigmatic game to serve as the setting of the discussion. By far the most discussed game is the *Prisoner's Dilemma*.[20] In the classic, two-person Prisoner's Dilemma, each player has a dominant strategy—a strategy

---

15. *See* Michel Rosenfeld, *Contract and Justice: The Relation Between Classical Contract Law and Social Contract Theory*, 70 IOWA L. REV. 769, 824, 847–78 (1985).

16. *See id.*

17. Larry Samuelson, *Evoloution and Game Theory*, 16 J. ECON. PERSP. 47, 47–48 (2002).

18. *Id.*

19. Brian Skyrms et al., *In the Light of Evolution VIII: Darwinian Thinking in the Social Sciences*, 111 PROC. NAT'L ACAD. SCI. 10781, 10783 (2014).

20. *See, e.g.*, DAVID GAUTHIER, MORALS BY AGREEMENT 79–82 (1986).

that pays off better no matter what the other does—and this strategy is essentially to opt out of the social contract. Everyone would be better with a social contract, but it is in each person's individual best interest to not join. Here is an example Prisoner's Dilemma[21]:

	Join the Contract	Opt Out
Join the Contract	3, 3	1, 4
Opt Out	4, 1	2, 2

A *public goods provision game* provides an example of a multiplayer game with a Prisoner's Dilemma structure. Individuals decide what to contribute to a public good. They keep what they do not contribute. They all share equally in the public good. If the value of the public good is increased by more than the value of a contribution causing the increase, then the higher the contributions the better for the group. If the increase caused by a $1 is less than $N, where N is the size of the group, then self-interest calls for free-riding.

Taking the Prisoner's Dilemma as paradigmatic of the problem of forming a social contract seems remarkably pessimistic, even for someone as pessimistic as a Hobbesian. A rational self-interested agent would not join in forming a social contract, and would defect from one that already existed. But Hobbes' project was to show that rational self-interested agents would support a social contract.

Evolutionary theory provides a different perspective. There is a population of individuals. What are strategies in classical game theory become types of individuals. If the individuals in the population meet at random to participate in the interaction that constitutes the game (and the population is large), then analysis is approximately in agreement with rational choice game theory. But if meeting is not independent of the type of the individual—biologists talk of *population structure*—then results can be quite different. Because we are interested in cultural evolution, we can ask whether there are cultural or social reasons why encounters may be non-random. Of course there are. There is often *social structure* that induces encounters that correlate with behaviors. We must take this social structure into account.

Now suppose that in a Prisoner's Dilemma interaction cooperators (social contractors) always meet each other, and thus so do defectors (opt outers). Then cooperators do better than defectors. *Social structure can change everything*! This is the first great lesson that evolutionary game theory has for the theory of the social contract.

21. The numbers in the cells are respectively row's payoff, column's payoff.

It is thus important for social contract theory to study endogenous sources of correlation, of social structure, that can affect the outcomes of interactions. All of the so-called solutions to the Prisoner's Dilemma that have been advanced in biology and political science—kin selection, group selection, reciprocity in repeated games, etc.—are all ways of inducing correlation. They are all ways of making it more likely than chance that cooperators meet cooperators and defectors meet defectors. If correlation were perfect then cooperators would take over and that would be the end of the story. But endogenous correlation devices may not deliver perfect correlation, and then the story becomes more complicated and more interesting.

Another way of looking at endogenous correlation in the Prisoner's Dilemma is to model it as a result of interactions in a bigger game—in a bigger social interaction, where interactions in the bigger game generate correlation in the embedded smaller game. Reciprocal altruism is a transparent example, where repeated interactions form the context of the bigger game, and reciprocators generate correlation in the embedded Prisoner's Dilemmas. If the bigger games "solved" the Prisoner's Dilemma, we would expect the outcome to always be cooperation, but it is not so. This is because they do not generate perfect correlation in the embedded Prisoner's Dilemmas. The bigger game often has the structure of a *Stag Hunt*, as I detail in my book *The Stag Hunt and the Evolution of Social Structure*.[22]

In the Stag Hunt game there are two equilibria, one where both players cooperate and one in which neither does. Here is an example of the Stag Hunt:

	Join the Contract	Opt Out
Join the Contract	4, 4	0, 3
Opt Out	3, 0	3, 3

Here if you are in the contract, you want to stay in. Opting out, if your partner is in, would be stupid. But if you are both out, opting in if your partner does not do so as well would also be stupid. The social contract and the state of nature are both equilibria. Remember Hume's rowboat? If we both row, it is the best outcome. If only one person rows, the boat goes around in circles and he would have been better off not rowing.[23] It is a Stag Hunt.

Just as there are multi-person Prisoner's Dilemma's there are multi-person Stag Hunts. *Public goods provision games with a threshold* provide

22. BRIAN SKYRMS, STAG HUNT, *supra* note 6, at 9–12.
23. HUME, *supra* note 2.

examples.[24] If cooperation falls below the threshold, then the whole enterprise fails. If you are next to the threshold, then a purely selfish individual would be better off cooperating and thus putting the group over the threshold to share in the fruits of cooperation than to hold back. But if you are well above the threshold, self-interest would call for free-riding.[25]

The Stag Hunt may fit Hobbes better than the Prisoner's Dilemma, as Curley remarks in his introduction to the *Leviathan*.[26] It may fit the major transitions of evolution better, as Maynard-Smith and Szathmary[27] remark at the end of their book on that subject. The problem of instituting the social contract is now one of getting from a bad equilibrium to a better one. At the bad equilibrium there is a risk in trying to get to the good one. It is possible to overcome the risk if it is possible to achieve more correlation, so that it is possible to trust those with whom one associates to do their part.

While some frame the social contract in terms of the Prisoner's Dilemma, or the Stag Hunt, others see it as better represented as a *Nash Bargaining* game.[28] The fruits of cooperation are a potential resource to be divided. If there is no agreement as to how to do this, the potential gains are not realized. Formally, the Nash bargaining game has the following structure: each player makes a final, bottom-line demand.[29] If the demands do not exceed the total resource then the players get what they demand.[30] If not they get the disagreement outcome.[31]

There is again a totally non-cooperative equilibrium in this game; each individual demands everything. And there are an infinite number of possible social contract equilibria, corresponding to the possible ways of splitting things up. Each individual may favor a contract that gives her more, but if they do not agree there will be nothing to divide. This brings a new problem, selection from the multiplicity of possible social contracts, to center stage.

The Hobbesian state of nature is a state of unrestricted resource competition. Evolutionary biologists' favorite game theory model of resource

---

24. *See* Jorge M. Pacheco et al., *Evolutionary Dynamics of Collective Action in N-Person Stag Hunt Dilemmas*, PROC. ROYAL SOC'Y B: BIOLOGICAL SCI. 315, 316 (2009).

25. *See id.*

26. Edwin Curley, *Introduction* to THOMAS HOBBES, LEVIATHAN vii, xxvii (Edwin Curley ed., 1994) (1651).

27. JOHN MAYNARD SMITH & EÖRS SZATHMARY, THE MAJOR TRANSITIONS IN EVOLUTION 261 (1997).

28. *See, e.g.*, BINMORE, JUST PLAYING, *supra* note 7, at 77. Binmore uses the theory of indefinitely repeated games to argue that "the game of life" has the structure of a bargaining game. *Id.* at 4–5.

29. SKYRMS, SOCIAL DYNAMICS, *supra* note 6, at 3–4.

30. *Id.* at 14–15.

31. *Id.*

competition has not yet been mentioned. It is the *Hawk-Dove* game (known in different circles as the game of "Chicken"). In contesting a resource Hawks fight to the end, Doves give up against Hawks and split the resource when they interact with each other. It is best to be a Hawk against a Dove. Since Hawks fight each other to the death, it is worst to be a Hawk against another Hawk.

	Hawk	Dove
Hawk	0, 0	4, 1
Dove	1, 4	2, 2

Evolutionary dynamics pushes towards a mixed population of Hawks and Doves. If there are lots of Doves, it is better to be a Hawk. If there are lots of Hawks, it is better to be a Dove.

So far there is no social contract in sight. But Maynard Smith and Price[32] suggest that we expand the model. Individuals who have been occupying a habitat can acts as owners, while if exploring a new habitat can act as intruders. Then they suggest that evolution may lead to a conditional strategy for resource competition: fight fiercely if owner; retreat if intruder. This strategy avoids the conflict inherent in a mixed population of pure Hawks and pure Doves. In technical terms it is a correlated equilibrium, a *correlated convention* in the sense of Vanderschraaf.[33]

Which of the foregoing is the correct model to use in a theory of the social contract? That is not a good question. They are all correct. They all capture aspects of a pluralistic social contract. And, as Hume reminds us, there are others: the enforceability of contracts, the development of a means of exchange, and the evolution of conventions of language.

With respect to language, evolutionary game theory gives us a start, in the form of signaling games. These games were introduced by David Lewis in *Convention* for the purpose of giving a game theory model of conventional meaning. We illustrate with the simplest example. Nature flips a coin to choose one of two states. The first player, the sender, observes the state and chooses one of two signals—say red or green—to send to the receiver. The receiver guesses the state and acts accordingly. In Lewis, sender and receiver have common interest in the receiver acting correctly. If the receiver guesses the state correctly, they both get paid; otherwise they do not.

32. John Maynard Smith & George Price, *The Logic of Animal Conflict*, 246 NATURE 15, 15–18 (1973).
33. Vanderschraff, *Correlated Equilibrium*, *supra* note 1.

	Receiver guesses 1	Receiver guesses 2
Nature chooses state 1	1, 1	0, 0
Nature chooses state 2	0, 0	1, 1

Sender and Receiver both want the information about the state to get across. They can achieve this by forming a social contract about the meaning of signals. There are two such contracts available. One says *red means state one; green means state two*. The other says *green means state one; red means state two.* Both are equally good. They are both equilibria. There are also states of nature in which the sender ignores the states in choosing whether to send red or green, and the receiver ignores the signal in choosing which state to guess. These are no-communication equilibria.

As always, we have the problem of getting from the state of nature to the social contract. But there is also a new problem. That is that there is complete symmetry between the two communication conventions. There is no possible reason for the individuals involved to choose to coordinate on one rather than the other. Of course, if they had language they could talk about it beforehand and agree. But we are trying to explain the emergence of language. We do not want to try to explain the emergence of language in a way that has to presuppose that we already have it.

But if we look at evolutionary dynamics, it turns out that dynamics breaks the symmetry, and meaningful signaling emerges spontaneously. Sometimes one convention emerges, sometimes the other. There are many details and variations,[34] but enough has been said to make the second great point that evolution has for game theory. *It is important to look at the dynamics*. This becomes all the more true as one looks more deeply into the subject.

III. PUTTING PIECES TOGETHER

Any particular contract assembles constituent conventions into a whole. Sometimes putting social interactions together produces a whole whose properties are not immediately evident from the properties of its parts. We start by putting together a signaling game with a Stag Hunt game. This is a nice example because we have analyses from the perspectives of both large population and small population evolutionary dynamics, and the analyses give the same qualitative conclusion. There are also similar applications of the same basic ideas to other games that we have discussed.

34. *See* SKYRMS, SIGNALS, *supra* note 6, at 145–48. And for an update, see Simon Huttegger et al., *Some Dynamics of Signaling Games*, 111 PROC. NAT'L ACAD. SCI. 10873, 10873 (2014).

Players exchange costless signals before playing a Stag Hunt game. As before, signals have no pre-existing meaning. If they are to acquire meaning, it must co-evolve with play in the Stag Hunt. Suppose that there are two signals available to each player, say red and green as before. Then a strategy in this extended game specifies: (i) what signal to send, (ii) what action to take if the other player sends red, and (iii) what action to take if the other player sends green. There are now eight strategies instead of two.

Consider the Stag Hunt play with and without signals in a large population, with random encounters and strategies being imitated (or replicated) with probability proportional to payoff received. If there are lots of Stag Hunters, it is best to hunt Stag, and evolution lets Stag hunting take over. If there are lots of Hare Hunters, then it is best to hunt Hare, and evolution lets Hare Hunters take over. The question here is one of relative size of basins of attraction. Starting with randomly chosen population proportions, how likely is it to end up with all Stag Hunters and how likely with all Hare Hunters? Brushing past details,[35] we jump to a qualitative answer. *Pre-play signaling dramatically increases the basin of attraction of Stag Hunting*, even though the signals have no pre-existing meaning.

How does it work for other games? Suppose, in the same setting, we consider a finite version of the Nash Bargaining game, where demands are restricted to one-third, one-half, or two-thirds of the resource being bargained over. To these three strategies are added three signals, yielding eighty-one strategies in the expanded game. Again, the pre-play signals have a dramatic effect on the size of basins of attraction, making it much more likely to end up with the equal split.

We can test the robustness of our qualitative conclusion regarding the effectiveness of signaling in the Stag Hunt game, by asking the same questions in a different evolutionary setting. We now consider a small finite population playing the Stag Hunt. Sometimes someone leaves the group. A new person enters and imitates an existing strategy in the group, chosen with probability proportional to the success of strategies. This process eventually leads to a uniform population, where everyone in the group has the same strategy. It remains uniform because newcomers only have one strategy to imitate.

We now modify this picture by supposing that once in a while, but rarely, a new person—an innovator, a mutant—enters the group with a strategy of the game not represented in the previously uniform group. This probabilistic dynamics never settles down to an equilibrium; there are always mutants

35. Details can be found in Brian Skyrms, *Signals, Evolution and the Explanatory Power of Transient Information*, 69 PHIL. SCI. 407, 407–08 (2002).

waiting in the wings. The question to ask here is what proportion of the time the group spends in a certain state. In our application of these dynamics we ask how much of the time does the group spend hunting Stag and how much hunting Hare? This can be calculated analytically.

As before, we compare the Stag Hunt game with signals to the game without signals. Although the question takes a different form here, our qualitative conclusion from the large population case survives. *Signaling favors Stag Hunting*. Furthermore, the present study allows us to say more. The more signals there are, the better for cooperation, the more time is spent Hunting Stag. With just four or five signals, the transformation of a game that originally favored non-cooperation to one that favors cooperation is quite remarkable.[36]

The same study has another surprise. Pre-play signaling even helps cooperation in the Prisoner's Dilemma.[37] Again, more signals are better. With six signals we can have a group spending most of its time cooperating. Mutants can use signals to set up correlation and cooperate; other mutants can fake these signals and victimize the cooperators; other mutants can use other signals to correlate and cooperate. Round and round it goes, and it all is a question of timing. With enough signals quite a bit of cooperation is possible.

For a different sort of example, we combine the Stag Hunt game with co-evolving *social structure*. Social structure is represented by an interaction network.

Individuals start out by interacting at random, but gradually learn to interact more with those with whom they have had good experiences. What generates these experiences? In this example, the interaction is playing the Stag Hunt game with the payoffs from that game driving the evolution of the network structure.

Payoffs may lead individuals to modify their social network, but they may also lead them to modify their strategies in the Stag Hunt game. Thus we have a co-evolution social structure and strategy. There is a structure modification dynamics and there is a strategy revision dynamics. The outcome will depend on the interaction of these two dynamic processes.

It turns out that their relative speeds make all the difference. With frozen social structure Stag Hunters may be locked into interactions with Hare Hunters. Then they give up attempts at cooperation and convert to Hare Hunting themselves. With rapidly evolving social structure, Stag Hunters

36. Francisco C. Santos, Jorge M. Pacheco & Brian Skyrms, *Co-evolution of Pre-play Signaling and Cooperation*, 274 J. THEORETICAL BIOLOGY 30, 33 fig.4 (2011).

37. *Id.* at 34 fig.6.

learn to associate with each other and prosper. Then, Hare Hunters gradually convert to Stag Hunting.[38]

There are lots of other ways to put simple games together to make more complex games. Signaling can take place on a signaling network, where the network has multiple senders and receivers for the nodes of the network.[39] Pre-play signaling for multiple senders and receivers can be combined with a multiplayer Stag Hunt to form a model of quorum signaling.[40]

I close this Section with one combination especially relevant to the framing of social norms.[41] Social norms can overlap, with multiple norms applying to the same situation, and giving conflicting results.[42] Multiplicity of applicable norms leads to sensitivity to framing effects.

This makes perfect sense from an evolutionary point of view. Norms evolve to govern classes of situations. These classes may overlap. Thus, a given situation may fall under a number of more or less general norms. Social psychologists know that experimental behavior can be manipulated by framing the same decision differently. We can interpret the framing language as a signal that triggers a relevant norm. For example, this can be used to explain apparently anomalous behavior in an Ultimatum Bargaining game.

In ultimatum bargaining, there is a resource to be divided as in Nash bargaining, but the players here have asymmetric roles.[43] One player, the proposer, gets to start by saying how much of the pie she wants.[44] The second player, the responder, can say yes, in which case the proposer gets what she asked for, and the responder gets the rest.[45] Or she can say no, in which case no one gets anything.[46]

The game is famous because laboratory experiments showed behavior that disagrees with what was then the conventional economic analysis. The

---

38. *See generally* Brian Skyrms & Robin Pemantle, *A Dynamic Model of Social Network Formation*, 97 PROC. NAT'L ACAD. SCI. 9340 (2000).

39. Brian Skyrms, *Evolution of Signaling Systems with Multiple Senders and Receivers*, 364 PHIL. TRANSACTIONS ROYAL SOC'Y B: BIOLOGICAL SCI. 771, 771 (2009).

40. There is a finite population analysis in Jorge M. Pacheco, Vitor V. Vasconcelos, Francisco C. Santos & Brian Skyrms, *Co-evolutionary Dynamics of Collective Action with Signaling for a Quorum*, PLOS (Feb. 23, 2015), http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004101.

41. Brian Skyrms & Kevin J.S. Zollman, *Evolutionary Considerations in the Framing of Social Norms*, 9 POL. PHIL. & ECON. 265, 265–66 (2010).

42. *Id.* at 266.

43. Werner Guth et al., *An Experimental Analysis of Ultimatum Bargaining*, 3 J. ECON. BEHAV. & ORG. 367, 370–71 (1982) (providing a basic overview of the Ulimatum Bargaining Games).

44. *Id.* at 371.

45. *Id.*

46. *Id.*

conventional analysis argued that if the receiver is offered only a pittance, she still will take it, because a pittance is better than nothing. The proposer knowing this, will offer the receiver only a pittance, and claim the lion's share for himself.

But in experiments, the proposer was usually not so greedy, and the responder would frequently decline derisory offers. There is by now an enormous literature, varying the experiment, explaining the result, challenging the results universality across different culture, and so forth. A lot of framing effects are documented in the experimental literature. This all is quite comprehensible if we remember that the norms being invoked by subjects did not evolve in an environment composed solely of ultimatum bargaining experiments.

There may be a norm about responding to an ultimatum, or a norm for bargaining in general. It is of some interest to look at the evolutionary dynamics arising from a mixture of ultimatum bargaining and Nash bargaining. The norm for an equal split can evolve more readily in this mixed environment than in either an environment of pure Nash bargaining or in one of pure Ultimatum Bargaining.[47] There are other, more global norms that could apply. Economists talk about a norm of "not leaving money on the table." Norms of fairness or equity have been invoked. Experimental economists usually try to avoid framing effects as much as possible. They may or may not even call this a bargaining game. This gives subjects great latitude in choosing the norms that apply. If one looks at the individual experimental data in all these experiments, instead of averages, one sees all sorts of different behavior. This is just what we should expect.

IV. TAKING THE MODELER OUT OF THE MODEL

The modeler chooses simple, easily analyzed games to model strategic interactions of interest. We considered putting simple games together to form more complex games. But so far, it has been the modeler putting the simple games together, and then applying evolutionary analysis. We would like the games themselves to evolve. We would like simple games to *self-assemble* to form more complicated games.

That is an entirely plausible request for evolutionary game theory to satisfy, but so far only a few baby steps have been taken in that direction. For signaling games, there is a model in when the players can *invent new signals*,

47. Kevin J.S. Zollman, *Explaining Fairness in Complex Environments*, 7 POL. PHIL. & ECON. 81, 94 (2007).

thus allowing the signaling game itself to evolve.[48] Invention is incorporated as part of a learning process. This idea could be applied more widely.

There are beginnings of a theory of self-assembling games.[49] In this theory, learning dynamics accomplishes what the modeler did in the last Section. Players can learn to compose modules; they can learn to put together simple games with which they are familiar to make more complicated games. And sometimes they can learn to use old game forms for new purposes, a process called *template transfer*.[50]

Of course this does not really take the modeler out of the model. These are just bigger models. But dynamic models, in which strategies, social structure, and the game itself co-evolve, are a step closer to the real world.

---

48. Jason McKenzie Alexander, Brian Skyrms & Sandy L. Zabell, *Inventing New Signals*, 2 DYNAMIC GAMES & APPLICATIONS 129, 129 (2012).

49. Jeffrey A. Barrett & Brian Skyrms, *Self-Assembling Games*, BRIT. J. PHIL. SCI., Sept. 13, 2015, at 1, 3.

50. *Id.*; Jeffrey A. Barrett, *The Evolution of Simple Rule Following*, 8 BIOLOGICAL THEORY 142, 147 (2013).