13

The Games People Play

PETER DESCIOLI
ROBERT KURZBAN

In the preface of *Evolution and the Theory of Games*, Maynard Smith (1982) observed that "game theory is more readily applied to evolutionary biology than to the field of economic behavior for which it was originally designed" (p. vii). This is largely because an important underlying assumption of game theory is that the agents it models are rational, making decisions based exclusively on the costs and benefits of available options. By the time Maynard Smith published his book, the fact that humans frequently departed from rationality in their decision making had been amply demonstrated, (e.g., Kahneman, Slovic, & Tversky, 1982), and research supporting this conclusion continues to accumulate.

As Maynard Smith (1982) pointed out, this criticism does not apply to genes. Genes can be considered to be agents that embody strategies that over the course of evolution are tested against alternative strategies. By the process of natural selection, genes (strategies) that lead to the best (fitness) outcomes with respect to others spread in a population. In this way, superior strategies emerge by virtue of the decision rules they employ. The fact that natural selection is sensitive only to fitness outcomes makes genes rational agents *par excellence*. Strategies are "chosen" by natural selection based solely on the reproductive outcomes of those strategies.

Some genes have been selected by virtue of their role in building cognitive mechanisms that compute responses to recurrent adaptive problems. When the adaptive problem involves strategic interaction, these mechanisms might function in a way that approximates game theoretic solutions. Crucially, this does not imply that these mechanisms will be well designed to solve game theoretic problems in the abstract, but only specific strategic problems. Thus, computational mechanisms designed for strategic interaction might or might not embody the principles of "rationality" that underpin neoclassical economics; they simply execute the computations that increased the fitness of genes relative to alternatives.

This idea, the intermediate role of cognitive mechanisms, justifies the preceding Maynard Smith quotation; humans should not be expected to be "rational" in the traditional sense of the term. However, the strategies embodied by human cognition can be informed by a consideration of the adaptive problems they were designed to solve. Considering these adaptive problems, especially interpersonal strategic interaction, can inform hypotheses about the mechanisms underlying the solution to these problems. Applying classical rational choice theory directly to human behavior embodies the same mistake as applying fitness maximization (Symons, 1992): It misses the crucial mediating role of cognition. Cognitive mechanisms were subjected to the rigors of natural selection, and this analysis should be used in theory construction (Cosmides & Tooby, 1994).

EXECUTING OR ALTERING EQUILIBRIUM STRATEGIES

A recurrent game theoretic structure can broadly select for two kinds of adaptations: (1) adaptations designed to execute equilibrium strategies, and (2) adaptations designed to alter the structure of the game to shift the equilibrium. The latter idea is the same concept that underlies contract theory in economics. When parties have an incentive to defect on a mutually agreed upon arrangement, binding contracts can change incentives such that the contractually compliant move is more advantageous than defection, because of the penalties for noncompliance. Contracts change payoffs and, hence, change the game structure.

The fact that adaptations can alter the game itself suggests that it is a mistake to examine games such as the prisoner's dilemma (PD) in isolation from neighboring[1] games. Trivers (1971) "solved" the PD by adding probabilistic repetition; when the probability of continuation is great enough, the interaction takes on the form of a stag hunt game, that is, it has two equilibria, one of which is mutual cooperation. Adaptations associated with reciprocal altruism, such as individual recognition and memory of interactions, might have been favored by natural selection, because they changed PD iterations into stag hunt game interactions. Because adaptations can change the game structure of interactions, it is crucially important to understand the relations among different games.

These considerations suggest that researchers interested in strategic interaction should first identify the game that best models the adaptive problem of interest. Then, game theoretic analysis can be employed to formulate hypotheses regarding cognitive solutions to the problem. Furthermore, researchers should carefully consider whether the "initial game conditions" can select for adaptations that act to change the game itself.

TOWARD A TAXONOMY OF GAMES

An important step in using game theory more effectively in psychology is to develop a taxonomy of games suited to that purpose. A good taxonomy should help identify the game that best models a given adaptive problem, and should clarify relations among games.

We start here by developing a basic taxonomy of the simplest type of strategic interaction: two organisms, each with two options. Previous taxonomies of 2 × 2 games have been developed (e.g., Rapoport, Guyer, & Gordon, 1976) but served a different, more exhaustive function. Even in the simplest of these, considering only ordinal payoffs (outcomes ranked 1–4), there are 78 "basic" strategically distinct games. Although useful to game theorists, this is a cumbersome guide to psychologists looking for the right game to model an adaptive problem.

Our approach requires a few basic concepts: dominance, equilibrium, and Pareto efficiency.[2] Briefly, one option (strictly) *dominates* another when

[1] By "neighboring" games, we refer to "neighborhood" in the mathematical sense. For example, a 2 × 2 game can be represented as an eight-element ordered set (the matrix payoffs), that is, an eight-tuple point, x, that belongs to the space R^8. A neighborhood of x, $N_\varepsilon(x)$, is the set of points (games) inside an eight-ball, B^8, with center x and radius $\varepsilon > 0$.

[2] For brevity, these concepts are simplified and applied only to 2 × 2 games. Regarding equilibrium, for this initial analysis, we ignore mixed strategy equilibria and consider only pure strategy equilibria.

the payoff of that option is better than the alternative, regardless of the other player's decision. An *equilibrium* is the case in which neither player can obtain a better outcome by changing strategy, assuming that the other player does not change strategy. Finally, *Pareto efficiency* (with two players) refers to an outcome such that there is no other outcome that makes at least one player better off while simultaneously making the other player no worse off.

Our analysis greatly reduces the number of simplest possible games. An ordinal 2 × 2 game can be reduced to a set of four strict preference relations ($\succ$) while retaining sufficient information to identify equilibria (see also Maynard Smith, 1982). One can think of this as ranking outcomes not only ordinally per se, but ordinally *conditional on the action of the other player.* For example, in the Pure Dominance game in Figure 13.1, if Player 2 chooses strategy *A2*, Player 1 prefers strategy *A1* to *B1*; if Player 2 chooses strategy *B2*, Player 1 prefers *A1* to *B1*. In symbols, for Player 1, $A1 \succ B1|A2$, and $A1 \succ B1|B2$. In this case, Player 1 has a *dominant strategy*, because $A1 \succ B1$ for both possible actions of Player 2. In contrast, in the Coordination game, Player 1's strategies exhibit no dominance. More generally, when preferences converge on a strategy, it is dominant, and when they differ depending on the action of the other player, strategies exhibit no dominance.

There are four possible configurations of two strict preference relations for each player, yielding 16 possible game matrices. Like Rapoport and colleagues (1976), we regard games with rows and/or columns and/or

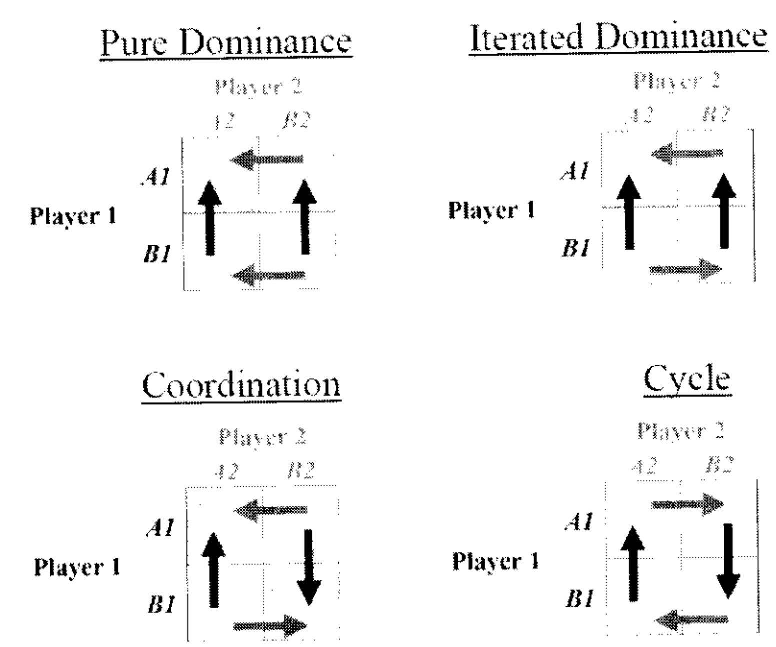

FIGURE 13.1. Four most basic games. Preference relations are represented as arrows pointing to more preferred outcome. Shaded outcomes are equilibria.

players interchanged as strategically equivalent. Thus, our analysis yields *four* strategically distinct games (see Figure 13.1).

A good taxonomy should distinguish among these four most basic games. Below we examine properties of each of these game types, including a partial set of subtypes, restricting discussion to those we believe to be the most relevant for evolutionary analysis.

The first game is a Pure Dominance game. Both players have strictly dominant strategies, resulting in a unique equilibrium. Two types of games can be distinguished by examining the Pareto efficiency of the equilibrium. When the equilibrium is Pareto deficient, the game is the familiar prisoner's dilemma. The adaptive significance of the PD is that the inefficiency generates a selection pressure for adaptations designed to alter equilibrium play: Adaptations associated with taking advantage of repeated PD games, discussed earlier, are examples. When the equilibrium is Pareto optimal, players benefit one another as a byproduct of doing what is best for themselves, essentially byproduct mutualism. The adaptive significance of the byproduct mutualism game is that it creates selection pressures favoring designs that are increasingly synergistic.

The second game is an Iterated Dominance game. Player 1 has a strictly dominant strategy. Player 2 does not have a dominant strategy, but does have a best response to Player 1's dominant strategy, yielding a unique equilibrium. Under these circumstances, the predictable behavior of Player 1 acts as a selection pressure on Player 2, favoring a design that makes use of Player 1's stable behavior. When the equilibrium (*A1*, *A2*) is better for Player 1 than (*A1*, *B2*), this will lead to selection for adaptations in Player 2 that complement Player 1's design, a process akin to mutualism. In contrast, when the equilibrium is a worse outcome than (*A1*, *B2*), Player 2 is selected for a kind of parasitism of Player 1.

The third game is a Coordination game. Neither player has a dominant strategy in this game, and there are two equilibria. Three subcategories of this game can be identified by comparing the equilibria. When the equilibria are identical, the game is a Pure Matching game. This situation should lead to adaptations designed to coordinate on one or the other equilibrium, possibly through signaling, again leading to mutualism. An important application of this game is Gil-White's (2001) analysis of cultural norms as solutions to coordination games.

When one equilibrium Pareto dominates the other, the game is a stag hunt (or assurance) game. This should similarly lead to adaptations designed to achieve the superior equilibrium, though organisms can get "stuck" in the inferior equilibrium due to path dependencies, design constraints, and so forth. When players have differing preferences with respect to the two equilibria, the game is a battle of the sexes (chicken, hawk–dove) game, where the mix of strategies will depend on the payoffs at each equilibrium.

The fourth game is a Cycle game. Neither player has a dominant strategy in this game, and there is no equilibrium. Two types of this game can be identified by examining the Pareto ordering of the outcomes. First, a cycle game in which no outcome Pareto dominates any other outcome amounts to a game of pure opposition, as is often characteristic of predator–prey interactions. This can lead to selection for adaptations designed to conceal likely future actions or randomize behavior (Miller, 1997).

Second, if at least one outcome is Pareto dominant to one other outcome, the game is an interesting mixed-motive game that can favor counterintuitive adaptations. Consider an organism that evolves to emit a costly signal that decreases its own payoffs to one strategy more than another. This can change the organism's preferences and, when communicated to the other player, can lead to reciprocal adaptations that allow a Pareto superior outcome to be obtained (see Figure 13.2; Zahavi & Zahavi, 1997). These weakness-is-strength adaptations are puzzling absent game theoretic analysis.

CONCLUSION

Game theory provides a set of useful tools for thinking about adaptations designed to negotiate recurrent strategic problems. Here we have provided

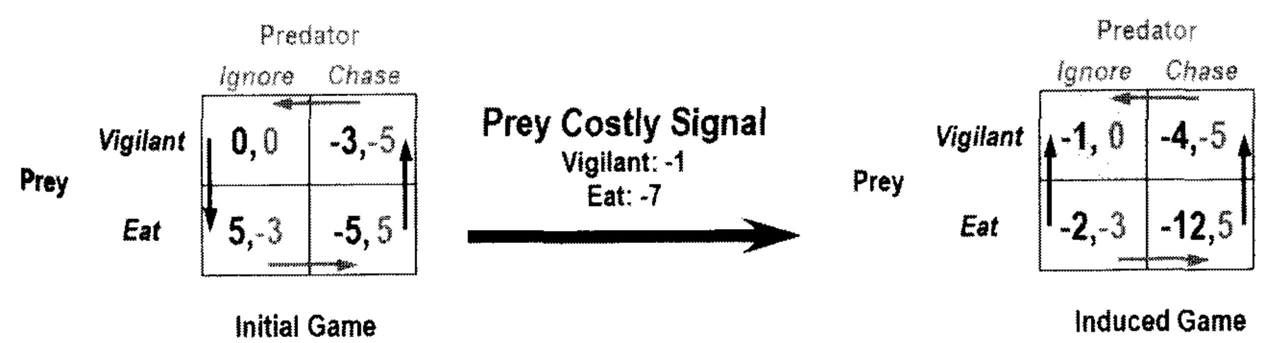

FIGURE 13.2. Prey signaling. Some predator–prey interactions exhibit Pareto ordering of outcomes, because both prefer no chase to a chase that results in escape, as in the Initial Game above. This Cycle game has a mixed strategy equilibrium at roughly [(.6 *Vigilant*, .4 *Eat*), (.3 *Ignore*, .7 *Chase*)] with expected payoffs of (−2.1, −1.2). Suppose that a signal (e.g., calling to predators) has a small cost to *Vigilant* Prey (who will be likely to escape if detected), but is very costly to *Eating* Prey (who will be less likely to escape if detected). The signal induces an Iterated Dominance game with Pareto superior payoffs (−1, 0); thus, design for both signaling and reception would be favored by selection. Calling to predators is one of a number of counterintuitive examples of predator–prey communication described by Zahavi and Zahavi (1997).

a simple taxonomy of strategic situations in the hope that this proves useful for developing hypotheses about design features associated with various domains of social interaction. By identifying the appropriate game, and by specifying relations with neighboring games, generating predictions about the adaptations designed to play the game in question should be possible. In summary, we hope this analysis facilitates a crucial task: characterizing as clearly and closely as possible the games people play.

REFERENCES

Cosmides, L., & Tooby, J. (1994). Better than rational: Evolutionary psychology and the invisible hand. *American Economic Review, 84,* 327–332.

Gil-White, F. (2001). Are ethnic groups biological "species" to the human brain? *Current Anthropology, 42,* 515–554.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Maynard Smith, J. (1982). *Evolution and the theory of games.* Cambridge, UK: Cambridge University Press.

Miller, G. F. (1997). Protean primates: The evolution of adaptive unpredictability in competition and courtship. In A. Whiten & R. W. Byrne (Eds.), *Machiavellian Intelligence II: Extensions and evaluations* (pp. 312–340). Cambridge, UK: Cambridge University Press.

Rapoport, A., Guyer, M. J., & Gordon, D. G. (1976). *The 2 × 2 game.* Ann Arbor: University of Michigan Press.

Symons, D. (1992). On the use and misuse of Darwinism in the study of human behavior. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 137–159). New York: Oxford University Press.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46,* 35–57.

Zahavi, A., & Zahavi, A. (1997). *The handicap principle: A missing piece of Darwin's puzzle.* Oxford, UK: Oxford University Press.

14

Dynamical Evolutionary Psychology and Mathematical Modeling

Quantifying the Implications of Qualitative Biases

DOUGLAS T. KENRICK
JILL M. SUNDIE

Although the topic of mathematical modeling draws a blank stare from many evolutionary psychologists, most of us are actually big fans, once we think about it. Trivers's classic arguments about reciprocal altruism, for example, or Haldane's kin selection quip about giving his life for two brothers or eight cousins, are simple models with which most of us are familiar. Mathematical modeling is merely a tool to extend logical reasoning, adding some numbers to increase precision.

Consider the classic prisoner's dilemma in Figure 14.1. In the standard setup on the left, each thief must decide whether to cooperate with his partner in crime (C), or defect on his partner by turning state's evidence (D). If B defects while A cooperates, B gets the best outcome (payoff of 7), and A the worst (payoff of 2). Though mutual cooperation yields the best group-level outcome, traditional economic models predict that each person will defect in the one-shot game.

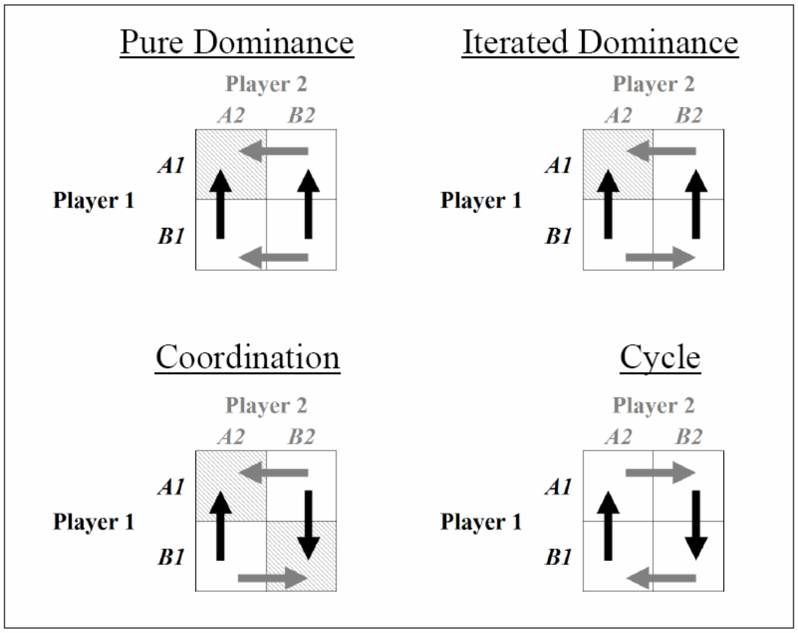

Figure 1. Four Most Basic Games. Preference relations represented as arrows pointing to more preferred outcome. Shaded outcomes are equilibria.

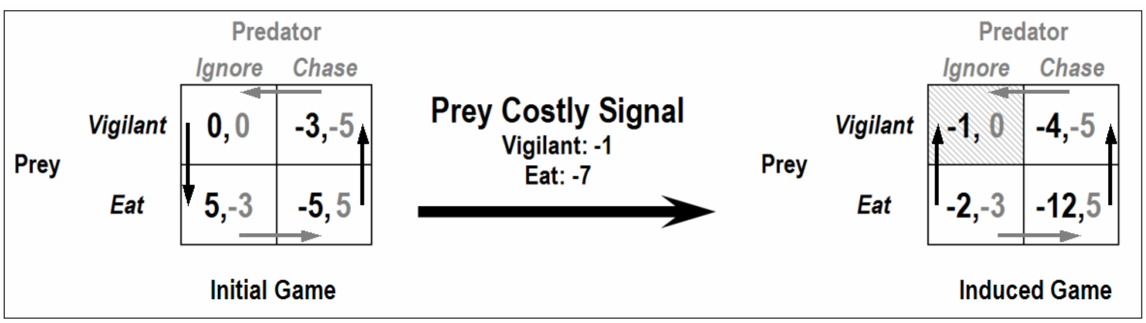

Figure 2. Prey Signaling. Some predator-prey interactions exhibit Pareto ordering of outcomes because both prefer no chase to a chase that results in escape, as in the Initial Game above. This Cycle game has a mixed strategy equilibrium at roughly [(.6 *Vigilant*, .4 *Eat*), (.3 *Ignore*, .7 *Chase*)] with expected payoffs of (-2.1, -1.2). Suppose that a signal (e.g., calling to predators) has a small cost to *Vigilant* Prey (who will be likely to escape if detected), but is very costly to *Eating* Prey (who will be less likely to escape if detected). The signal induces an Iterated Dominance game with Pareto superior payoffs (-1, 0) and, thus, design for both signaling and reception would be favored by selection. Calling to predators is one of a number of counter-intuitive examples of predator-prey communication described by Zahavi and Zahavi (1997).