P

Prisoner's Dilemma

Natalie A. Bloomston and Jonathan F. Prather
Program in Neuroscience and Department of Zoology and Physiology, University of Wyoming, Laramie, WY, USA

Overview

Game theory is a set of ideas developed with the goal of applying analytical rigor to understand decision-making. It includes consideration of key aspects that affect a decision, such as the nature of possible outcomes and the associated incentives (e.g., various degrees of payoff or punishment associated with outcome A versus outcome B). One of the most interesting aspects of game theory is that multiple participants are involved. Each participant can only make their individual choice, but the nature of the global outcome depends on both of their choices. In that way, each participant plays a role in shaping the global outcome, but neither has exclusive control over what happens. Each participant is assumed to be rational and acting in their own best interest, yet an outcome can sometimes emerge in which neither player achieves their optimal outcome. Thus, each participant chooses what is best for them, but that may not always yield the outcome that is best for the group. One of the most well-known ways in which this concept has been framed is the Prisoner's Dilemma in which two prisoners accused of the same crime are interrogated separately and asked whether or not they will confess to the crime. There are costs and benefits associated with their possible choices, and experts have used that scenario as a way to investigate the logical basis of decision-making. Specifically, how do individuals make decisions when risk and reward are involved, and what happens when the outcome that is best for the group is not what is best for each individual? In the following sections, we describe the history and concepts of the Prisoner's Dilemma, and we discuss ways in which ideas emerging from that analysis have been adopted by organizations such as companies and nations seeking to benefit in their interactions with partners and competitors.

Historical Background

The foundations of game theory include an assumption of rationality (Davis 1983; Leyton-Brown and Shoham 2008; Myerson 1991). Players in games are assumed to act rationally, meaning that they are expected to choose their actions based on logic and reasoning (Rapoport 1992; Skyrms 1990). This ideal has had to be adapted in the course of research into how mathematical models may account for human behavior, as the behavior of humans and other organisms can be very difficult to characterize and often deviates from a single ideal outcome

J. Vonk, T. K. Shackelford (eds.), *Encyclopedia of Animal Cognition and Behavior*,
https://doi.org/10.1007/978-3-319-47829-6_1404-1

(Glimcher et al. 2009; Thaler and Sunstein 2008). This is because different choices can each be considered rational despite the fact that they result in very different outcomes. For example, consider a case in which each of the members of a group is dependent on a shared resource in order to survive. If we consider the well-being of the group, we would expect that each member would act in such a way that they would not only benefit from the resource but also preserve it for continued shared use in the future. However, if we consider the well-being of each individual separately, we would expect each member to want to take more for themselves even though that would come at the expense of the group. In each of these possibilities, the members of the group are making rational choices, but the difference lies in whether they are focused on the well-being of the group or themselves as individuals. In this scenario, each individual is incentivized to make choices in a way that is less than optimal for the group. Each individual seeks to protect themselves at the expense of the others, and as a result both end up in a state that is worse than if they had worked together to coordinate their decision-making.

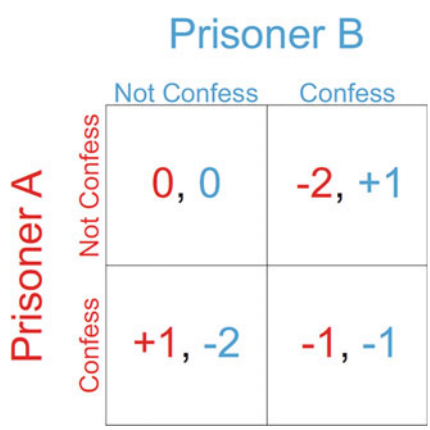
This paradox of decision-making is one of the most well-known ideas associated with game theory, and it lies at the heart of the Prisoner's Dilemma. These differences between expected and observed behaviors have intrigued experts in economics, psychology, and policymaking, and those researchers have sought to formalize their thinking about that process. In the first consideration of this process, researchers Merrill Flood and Melvin Dresher at the RAND Corporation envisioned a "game" in which two individuals each contributed to choosing an outcome that would affect them both (Flood 1958). Each player was faced with deciding what decision-making process would be best for them individually when the consequences were predictable but the behavior of the other player was not. This challenge of how to make the best decision in a mutually dependent outcome is the essence of what eventually came to be known as the Prisoner's Dilemma, but it awaited contributions from others before it was framed in the way that we consider it today.

Developing a Formalized Understanding of the Prisoner's Dilemma

When Flood and Dresher first proposed their hypothetical scenario, they and others quickly realized that it is an excellent example of challenges faced by participants in many forms of interaction, from board games to economics to international policymaking. A fundamental advance in the thinking about these scenarios occurred when Princeton mathematician Albert Tucker proposed the context of two people who were each accused of the same crime (Tucker 1950) (Fig. 1). Each person was seeking to minimize their punishment and maximize their benefit, but cooperation was not possible because each was being interrogated separately. In that scenario, each prisoner was told that:

(1) If one confesses and the other does not, the former will be given a reward of one unit (+1 in the value of the outcome) and the latter will be fined two units (−2 in the value of the outcome).
(2) If both confess, each will receive a lesser fine (−1 in the value of the outcome for each).

At the same time, each prisoner has good reason to believe that:



Prisoner's Dilemma, Fig. 1 The benefits (positive numbers), punishments (negative numbers), and possibility of going free (zero values) inherent in the decisions faced by each participant in the Prisoner's Dilemma

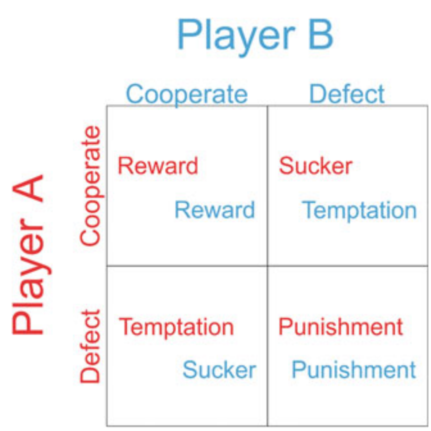
(3) If neither confesses, both will go free (0 value of the outcome for each).

In subsequent considerations of this scenario, it was further stipulated that both prisoners fully understand the nature of their choices and the associated consequences, neither prisoner has an opportunity for retribution or reward outside of the game, and neither prisoner has any loyalty to the other (Poundstone 1992). In other words, each prisoner acts in their own self-interest and makes their decision independently of the other.

In this scenario, the overall most beneficial result would occur if both players choose to not confess, as both would go free and neither would be fined. Thus, choosing to not confess could yield freedom, but it also exposes each decision-maker to the greatest possible risk if the other prisoner confesses. Therein lies the dilemma – should the prisoner choose the option that could enable them to achieve the best global benefit or the option that could enable them to avoid the worst individual punishment? Should they act in the best interest of the group or of themselves? If the prisoners were able to coordinate their strategies and form a cooperative solution, the safest and most beneficial option would be for them to each choose to not confess. However, in Tucker's scenario, the participants are not able to work together and form such a coalition. Therefore, each participant lacks the assurance associated with coordinated decision-making.

In the absence of coordinated assurance and collective action, each participant is incentivized to act in their own self-interest. In the case of the prisoners in this scenario, each is incentivized to confess, making confession the dominant strategy for both players. This dominance emerges because neither player is able to improve their individual payoff by altering their strategy (Rapoport 1992). For example, if Player A chooses to confess, Player B is better served to confess (−1 outcome) than to not confess (−2 outcome). Similarly, if Player A chooses to not confess, Player B is still better served to confess (+1 outcome) than to not confess (0 outcome). Regardless of what the other player does, confessing always results in the better payoff. An action profile such as this, where no single player can increase their payoff by deviating from their initial strategy while the other player remains unchanged, is known as Nash equilibrium (Nash 1950). Princeton mathematician John Nash recognized and described this pattern of making choices in the face of uncertainty, and the Nash equilibrium is considered the most rational outcome. In the scenario involving Players A and B, the Nash equilibrium is mutual confession (the bottom-right quadrant in Fig. 1). It results in a small fine, but it also enables them to ensure that they avoid the possibility of a much larger fine. Even though mutual cooperation (both choosing to not confess) would yield the best outcome for both, two players seeking to maximize their own payoff and minimize their own risk would be expected to choose the Nash equilibrium of mutual confession.

The rewards and punishments that drive the emergence of this behavior can be even more easily observed when the game is removed from the familiar context of prisoners and confessions to a more general context in which there are two players with the two options of "cooperate" or "defect" (Fig. 2). If Player A and Player B both cooperate, they will each receive the "reward" payoff (R). If they both defect, they will each receive the "punishment" payoff (P). If either player is the sole defector, they will receive the



Prisoner's Dilemma, Fig. 2 – The rewards and punishments associated with each player's decision in a generalized form of the Prisoner's Dilemma (adapted from Fig. 1 in Axelrod and Hamilton 1981) "temptation" payoff (T), while the other player who chose to cooperate when their opponent chose to defect will receive "sucker's" payoff (S). The magnitudes of the rewards or punishments may vary in different scenarios, but a condition like the Prisoner's Dilemma emerges when participants are asked to each engage in an independent choice and the relationships among those outcomes satisfy the condition that T > R > P > S (Axelrod and Hamilton 1981; Poundstone 1992; Rapoport et al. 1965).

Iterated Prisoner's Dilemma and the Emergence of Strategies

In the original vision of the Prisoner's Dilemma, the game was only played once (Poundstone 1992). In that context, each player must make their decision based only on the set of actions available to them and the payoffs or consequences associated with each action. Each player must seek the most beneficial outcome even in the face of a completely unknown choice to be made by the other player. This scenario has led to the emergence of insights about how to make rational decisions in the face of uncertainty, but it fails to allow for the possibility that interactions and the associated choices may occur more than once. When the game is played more than once, players can learn from the previous actions of their opponent and modify their own strategy accordingly. Therefore, the factors influencing each player's decision are not only the payoffs for each action but also their experience and knowledge of their opponent's past behavior as well as their intuition about their opponent's future behavior. This scenario in which each player makes a choice but repeated interactions are possible is referred to as the iterated form of the Prisoner's Dilemma (Poundstone 1992). When the game is played more than once, risks and rewards must be weighed not only in the immediate next step but also in light of past experiences and possible future interactions.

The opportunity for each player to learn from experience and modify their behavior during repeated interactions gives rise to the possibility that each player may develop different types of strategy (Axelrod and Hamilton 1981). In the relatively simple game envisioned in the Prisoner's Dilemma (Fig. 2), both players have two possible actions: they can either cooperate or defect. Strategies can be broadly categorized as those that begin with cooperation (the so-called "nice" strategies) and those that begin with defection ("mean" strategies). When both players are nice, both get an immediate reward. If both players are mean, both players receive the punishment result. The incentive is greatest for a player to be mean when their opponent is nice (temptation payoff goes to the mean player, and the nice player receives the sucker's payoff). The names for each of these strategies reflect the outcome associated with the first choice that each player makes, but much more complex strategies can emerge when we consider how those outcomes may affect subsequent choices.

A player may change their responses based on the opponent's previous actions (a reactionary strategy). For example, a player may change away from cooperation only after their opponent chooses to defect for the first time. If that player remains in defection, it can be called a "Friedman" or "grudger" strategy because the player's behavior changed once in response to a mean action and remained fixed thereafter (Axelrod 1980a). Alternatively, a strategy may include an element of "forgiveness" in which behavior may continue to remain variable even after the opponent engages in a mean action. If short-term history dependence is incorporated into that strategy, then a "tit for tat" strategy can emerge in which a player changes their behavior to match that of the previous action made by the other player (Axelrod 1980a). In other forms of strategy, a player may change their response based on some inherent feature of that strategy in which changes are made preemptively in an attempt to gain advantage or to test the reaction of the opponent (a preemptive or testing strategy). For example, if a player detects that their opponent very regularly chooses to cooperate, then the best choice for that player would be to defect, as the player would receive the temptation payoff and the opponent would be stuck with the sucker's payoff. These types of contingent or preemptive strategies seek to probe the patterns of responses expressed by their opponent and to use that information to exploit other simple strategies to gain advantage.

In seeking to understand which strategies may be most beneficial when games are played more than once, researchers created different kinds of strategies and tested their effectiveness when pitted against one another in simulated competitions. In 1979, a researcher named Robert Axelrod held a simulated tournament in which 14 different types of iterated Prisoner's Dilemma strategies, as well as 1 random strategy, competed in an attempt to determine which types of strategies were most broadly beneficial (Axelrod 1980a, 1984). In this competition, the iterated version of the Prisoner's Dilemma still included two possible actions for each player. Those actions were still called "cooperate" and "defect," but the associated outcomes were considered in the context of earning points rather than avoiding a prison sentence (Axelrod 1980a). If both players defected, then they each gained a small amount of points. On the other hand, if both players cooperated, both gained a significantly greater amount of points. If one player defected when the opponent cooperated, then the defector received the maximum amount of points, and the player that cooperated received zero points. The value of any strategy relative to others is ultimately dependent on the magnitudes of those point values in the payoff matrix, but this arrangement of temptation > reward > punishment > sucker's payoff is the same as described for the generalized form of the Prisoner's Dilemma.

In Axelrod's tournament, strategies that began with cooperation ("nice" strategies) tended to fare better than the "mean" strategies that began with defection (Axelrod 1980a). The winning strategy in that competition was the "tit for tat" strategy, created and entered by Anatol Rapoport, in which the player cooperates in the first round and then in each subsequent round simply mimics what the opponent selected in each previous round (Kopelman 2020). In this way, the opponent receives the treatment that they gave in the immediately preceding round, hence the name tit for tat. This strategy is one in which the player retaliates when provoked but otherwise cooperates. The success of this strategy against a wide variety of opponents suggests that there is an advantage in not only behavioral flexibility but also adaptation in the patterning of those behaviors. A similar tournament followed up on that first tournament, this time with an even greater variety of 62 different strategies and 1 random strategy pitted against one another (Axelrod 1980b). Despite the greater variety of opponents, tit for tat again emerged as a broadly beneficial basic strategy. Although tit for tat is not always the overall winner in these types of tournaments, it tends to fare very well across repeated interactions with many different types of strategies, further confirming the value of behavioral flexibility and history-dependent adaptation. A topic of ongoing research is the degree to which different depths of history dependence may be beneficial. For example, what if a player alters their strategy based on not only the previous action of the opponent but also the nature of the collective outcome that it produced (the so-called "win-stay, lose-switch" strategy) (Imhof et al. 2007; Nowak and Sigmund 1993)? What if a player alters their behavior based on not only the most recent action by an opponent but also in response to other actions from the more distant past? In that light, researchers are exploring the degree to which it may be beneficial for strategic algorithms to be "forgetful," such that only the most recent events matter, or to have some sort of "memory", in which actions from the more distant past are influential in driving current actions.

Relevance of the Insights from the Prisoner's Dilemma for Real-World Applications

The scenarios and outcomes illustrated by the many forms of the Prisoner's Dilemma reveal the intricacies of decision-making that can arise in those fascinating contexts, but they leave open the question of how insights from the Prisoner's Dilemma may be applicable to decision-making in more practical terms. There are many types of interactions in nature and human society in which choices and consequences can be described using a payoff matrix very similar to those found in the

Prisoner's Dilemma (Axelrod and Hamilton 1981). In that way, the choices that participants must make in those contexts can be modeled as a form of the Prisoner's Dilemma, and insights from those scenarios can be applied to develop a better understanding of ourselves and processes in the world around us.

The iterated Prisoner's Dilemma has been used to model interactions in which trust between two participants is essential for their collective success. In the variation known as Rousseau's Stag Hunt, the familiar context of the Prisoner's Dilemma is modified such that the participants are two hunters (Skyrms 2004). Each hunter can pursue either stag or rabbit. The stag would provide a large meal that could feed all of the members of their group, but it is very challenging to hunt and requires the cooperation of both hunters in order to collect it. In contrast, a rabbit is a much smaller meal, sufficient to feed only the hunter, and is much easier to catch, requiring only one person to collect it. In this scenario, the greater reward of catching a stag requires the cooperative efforts of both hunters, and their success is uncertain. A hunter who chooses to pursue stag runs the risk of their partner choosing to not cooperate, and such a result would leave them with no food. A hunter who chooses to pursue rabbit has a much greater chance of meeting their own needs, but that reduction of risk means that they necessarily forgo the chance to collect the greater reward. The dilemma in this scenario emerges when rational players are drawn toward one option by the possibility of mutual benefit but are pulled towards the other option by the prospect of avoiding individual risk (Skyrms 2004). The most beneficial choice for the group would be to hunt stag, but that cooperative benefit would require cooperative action, and that would require each hunter to trust their partner. In that way, Rousseau's Stag Hunt models situations in which cooperation is difficult but still possible to achieve. Cooperation emerges through communication and experience as a form of social contract that is contingent on each person's belief about what the other person will do. With the accumulation of experience and the development of trust, it becomes more likely that each player can predict the behavior of the other, and they become more comfortable choosing the relatively risky option of cooperation. Thus, insights from the Prisoner's Dilemma reveal how accumulation of knowledge can provide benefits through not only prevailing in adversarial situations but also gradually learning how to work together in cooperative situations (Axelrod and Hamilton 1981; Axelrod 1984; Axelrod and Dion 1988).

The Prisoner's Dilemma can also be used to model the behavior of nations engaged in competition for superiority and security (Jervis 1978). For example, a nation that arms itself is prepared to prevail over an opponent or deter that opponent from attacking, but building those armaments comes at the cost of resources and time. Choosing to disarm would save costs, but that would also make that nation vulnerable to attack. Mutual development of arms by two nations could provide an effective deterrent of conflict, but both groups would encounter the costs of developing those weapons. Mutual disarmament would reduce the threat that each nation poses to the other, and it would also provide the additional benefit of reducing costs. As in the scenario of the two hunters that were dependent on one another for their well-being, trust is the central factor in this scenario. The least costly outcome would be mutual disarmament, but that would require trust of not only the other nation but also of other nations that may also be able to affect their well-being. Just as it was for the hunters, trust is a rare and precious commodity in any interaction between participants. Lack of trust makes self-interested behavior the only rational choice, and those types of interactions are evident in the relationships between many nations and other groups.

It is tempting to apply the lessons of the Prisoner's Dilemma across a wide range of practical scenarios to devise ways to incentivize specific behaviors or develop ways for participants to resolve conflict and perhaps even cooperate (Rapoport 1962). However, a fundamental challenge to such attempts often lies in the nature of the payoff matrix. In the Prisoner's Dilemma, the rewards or punishments associated with each choice are clear – if the prisoner chooses to cooperate, then there are only two actions that can possibly be performed by the other participant, and only one of two clearly defined outcomes can emerge. In the context of many attempts to apply the Prisoner's Dilemma to practical scenarios, those potential outcomes are not fully known and thus not possible to clearly define. In this way, not all scenarios are amenable to analysis through direct application of the Prisoner's Dilemma, especially those in which there is significant uncertainty regarding possible outcomes.

As evident in the scenario of the Stag Hunt, making rational choices based on the greater good requires trust between participants. This challenge is especially evident in the context of global climate change. As rising levels of atmospheric carbon dioxide appear to be driving correlated changes in climate, those changes are causing instability and undesired events that will continue if they are not countered by initiatives meant to reduce levels of CO_2 in the atmosphere (Raper and Braithwaite 2006; Solomon et al. 2009). All nations would benefit from reduced levels of CO_2, but the changes necessary to cause those reductions may be very challenging and costly (Reid et al. 2010). Therefore, each nation is simultaneously incentivized to maintain current behavior or otherwise avoid the expense associated with changing emission behaviors. If trust could be forged between two nations, then they could cooperate in their approach to this challenge, but without trust among all participants, those two cooperators would still be subject to defection and the sucker's payoff if other nations did not join them in their efforts. This challenge further illustrates how valuable trust between participants can be as a means of enabling coordinated action and rational choices that lead to the most broadly beneficial outcomes.

In many real-world scenarios such as each nation's decision of how to take action in response to climate change, there are many more than just two players. The challenge of finding ways to act in the best interest of the group becomes even greater as more players become involved. These types of scenarios in which many individuals share a resource and act independently in their own self-interest are sometimes called the Tragedy of the Commons (Hardin 1968). In those cases, individual users of a resource are incentivized to act in their own interest, but this comes at the expense of the common good. Through their actions, each user experiences short-term benefits, but the resource is adversely affected through their collective actions. Cases such as this in which there are many individual participants all incentivized to act in their own self-interest are especially difficult to analyze and even more difficult to devise solutions that lead to sustainable group well-being. Despite the scale of these daunting challenges, researchers continue to use the ideas inherent in the Prisoner's Dilemma as a context in which to seek generalizable insights into how the members of a population may achieve cooperation and thereby take collective action to provide collective benefits.

Cross-Reference

▶ Coordination Game
▶ Costs Benefits Analysis
▶ Decision-Making
▶ Dictator Game
▶ Game Theory
▶ Goal-Directed Behavior
▶ Nash Equilibria
▶ Problem-Solving
▶ Public Goods
▶ Punishment
▶ Rationality
▶ Stag/Hunt Game
▶ Traveler's Dilemma
▶ Trust Game
▶ Ultimatum Game
▶ Win-Shift
▶ Win-Stay

References

Axelrod, R. (1980a). Effective choice in the Prisoner's dilemma. *Journal of Conflict Resolution, 24*(1), 3–25. https://doi.org/10.1177/002200278002400101.

Axelrod, R. (1980b). More effective choice in the Prisoner's dilemma. *Journal of Conflict Resolution, 24*(3), 379–403. https://doi.org/10.1177/002200278002400301.

Axelrod, R. (1984). *The evolution of cooperation*. New York City, NY: Basic Books.

Axelrod, R., & Dion, D. (1988). The further evolution of cooperation. *Science, 242*(4884), 1385–1390.

Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. *Science, 211*(4489), 1390–1396.

Davis, M. (1983). *Game theory: A nontechnical introduction*. Mineola, NY: Dover Publications.

Flood, M. M. (1958). Some experimental games. *Management Science, 5*(1), 5–26. https://doi.org/10.1287/mnsc.5.1.5.

Glimcher, P., Camerer, C., Fehr, E., & Poldrack, R. (2009). *Neuroeconomics: Decision making and the brain*. Cambridge, MA: Elsevier Academic Press.

Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 1243–1248.

Imhof, L. A., Fudenberg, D., & Nowak, M. A. (2007). Tit-for-tat or win-stay, lose-shift? *Journal of Theoretical Biology, 247*(3), 574–580. https://doi.org/10.1016/j.jtbi.2007.03.027.

Jervis, R. (1978). Cooperation under the security dilemma. *World Politics, 30*(2), 167–214.

Kopelman, S. (2020). Tit for tat and beyond: The legendary work of Anatol Rapoport. *Negotiation and Conflict Management Research, 13*(1), 60–84. https://doi.org/10.1111/ncmr.12172.

Leyton-Brown, K., & Shoham, Y. (2008). Essentials of game theory: A concise multidisciplinary introduction. In *Synthesis lectures on artificial intelligence and machine learning* (1st ed., Vol. 2, Issue 1). Morgan & Claypool Publishers LLC. https://doi.org/10.2200/s00108ed1v01y200802aim003.

Myerson, R. B. (1991). *Game theory: Analysis of conflict*. Cambridge, MA: Harvard University Press.

Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences, 36*(1), 48–49.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's dilemma game. *Nature, 364*(6432), 56–58. https://doi.org/10.1038/364056a0.

Poundstone, W. (1992). *Prisoner's dilemma*. New York City, NY: Doubleday Books.

Raper, S. C. B., & Braithwaite, R. J. (2006). Low sea level rise projections from mountain glaciers and icecaps under global warming. *Nature, 439*(7074), 311–313. https://doi.org/10.1038/nature04448.

Rapoport, A. (1962). The use and misuse of game theory. *Scientific American, 207*(6), 108–119. https://www.jstor.org/stable/24936389?seq=1#metadata_info_tab_contents.

Rapoport, A. (1992). Game theory defined. *Rationality and Society, 4*(1), 74–82. https://doi.org/10.1177/1043463192004001009.

Rapoport, A., Chammah, A. M., & Orwant, C. J. (1965). *Prisoner's Dilemma: A study in conflict and cooperation* (Vol. 165). Ann Arbor, MI: University of Michigan Press.

Reid, W. V., Chen, D., Goldfarb, L., Hackmann, H., Lee, Y. T., Mokhele, K., Ostrom, E., Raivio, K., Rockström, J., Schellnhuber, H. J., & Whyte, A. (2010). Earth system science for global sustainability: Grand challenges. *Science, 330*(6006), 916–917. https://doi.org/10.1126/science.1196263.

Skyrms, B. (1990). *The dynamics of rational deliberation*. Cambridge, MA: Harvard University Press.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge, United Kingdom: Cambridge University Press.

Solomon, S., Plattner, G. K., Knutti, R., & Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences of the United States of America, 106*(6), 1704–1709. https://doi.org/10.1073/pnas.0812721106.

Thaler, R., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Tucker, A. (1950). A two-person dilemma (unpublished notes). In E. Rasmusen (Ed.), *Readings in games and information* (pp. 7–8). Blackwell Publishers: Oxford.