JARGON ALERT

Prisoner's Dilemma

BY DANIEL BROOKS

Rick and Kyle are two aspiring thieves with plans to get rich from reselling antiques on the black market. For their first heist, they set out to rob an antique shop in the countryside. They break into the shop's back door. When the alarm is triggered, they head to their getaway car. As they flee, a state trooper pulls the two over for speeding. Noticing that the car matches the eyewitness description of the one fleeing the scene of the crime, the trooper searches the car. He finds two unregistered firearms, and the men are locked up in the county jail.

Rick and Kyle are placed into separate cells. No communication is allowed between them. The officer goes to each cell and gives them the same options: Confess to the robbery or stay silent. If one confesses and the other stays silent, the confessor will receive no jail time, but the other must serve five years. If both stay silent, each gets one year in jail for the firearms charges. And if both confess, they each receive three years in jail.

Assuming that Rick and Kyle are self-interested, each will confess. Rick is worried that Kyle will not stay silent. So, if Rick stays silent, he gets five years, but if he confesses he gets three years. Even if Kyle decides to remain silent, Rick would still confess because no jail time is better than one year in jail. (Similar thoughts are running through Kyle's mind.)

The dilemma is that, for each suspect, confessing is the better choice no matter what the other person does. But, as a whole, they are worse off because they end up with a total of six years in jail when they could have received a total of two years if they both stayed silent.

Countless variations of this "prisoner's dilemma" story have been pondered ever since mathematician Albert W. Tucker first coined the term and formalized the game in 1950. Yet the punch line remains the same: Rational individuals acting in their own interest can result in suboptimal outcomes in the aggregate.

A common application of the prisoner's dilemma intuition is in the analysis of the conflict inherent in individual and group decisions. Members of a group that act in their own self-interest can end up making the group worse off than if everyone were to cooperate.

Examples may be found in the real world. Assume, for instance, that greenhouse gas emissions are responsible for global warming and that, all else equal, it would be desirable to curb climate change. Nations face a choice to either reduce greenhouse gas emissions or maintain the status quo. If enough nations cooperate, emissions will fall and the temperature stabilize. But since reductions in emissions require costly actions, any nation could sit on the sidelines and let other nations bear that cost while they enjoy the benefit. With enough selfish nations on the bench, cooperation breaks down and everyone loses relative to what they could achieve if they worked together. This might also be recognized as a "free-rider" problem. Prisoner's dilemmas can be seen in this light as an example of such a problem.

The logic of the prisoner's dilemma makes a big assumption about individuals: It presumes that they care mainly about their self-interest. Economic experiments have tested this assumption. When looking at a variety of prisoner's dilemma experiments from 1958 to 1992, Dartmouth College economist David Sally found that when participants played the games, on average, they tended to act selfishly only about half the time. A plausible explanation for this is that the players are less prone to selfish behavior than economists predict.

Another explanation might be that cooperation is actually an optimal strategy in the real world where people interact with each other repeatedly over time. To test this, Robert Axelrod, a social scientist at the University of Michigan, organized a tournament. Academic colleagues were invited to submit a computer strategy, which was to be repeated a number of times.

As it turns out, the exclusively selfish strategies did very poorly. The one that fared the best was also the simplest: the "tit for tat" strategy developed by a mathematical psychologist. It required the player to cooperate on the first move and then choose the same strategy that the opposing player picked on the previous turn. While the strategy gives the benefit of the doubt to the opposing player, it also lets him know that a lack of cooperation will not go unanswered. The enforcement of this implied social norm and the nature of reciprocal behavior over repeated rounds of the game might explain the rate of cooperation in a variety of experiments.

Or, to put it another way, even inherently selfish individuals may tend to cooperate more often when benefits to cooperation over the long run outweigh the benefits over the short run. It is through these experiments that economists have been able to mine a wealth of economic and social insight that arises from the hypothetical predicament of two prisoners. **RF**

ILLUSTRATION: TIMOTHY COOK