Trust and Cooperation: Analysis and Critique of the Rational Actor Model

Russell Hardin writes that trust is key to cooperation. For example, trust is an important part of an iterated prisoner's dilemma: unless the prisoner's trust one another, it will be difficult to keep choosing to cooperate. Unless trust is first established, the fear of defection is too high. Moreover, to trust is not a strategic act. An actor either trusts or distrusts, and people do not choose to trust: "I do not trust you in order to gain from interacting with you. But because I do trust you, I can expect to gain from interacting with you if a relevant opportunity arises (516)." But is it really necessary to introduce the notion of trust in order for cooperation to result (if we assume that actors are interest-maximizing egoists)? I will first display cases where this seems not to be the case: here I will focus on Robert Axelrod and Michael Hechter's work. Nonetheless, I will argue that in cases where cooperation occurs in environments with no material benefits being produced, a rational choice model that dispenses with normative explanations (such as one which incorporates the notion of trust) is limiting.

What is perhaps ingenious about Robert Axelrod's method is that he models interaction of players that have no option but to act as rational interest-maximizing egoists, for his players are computer programs. Computer programs do not have values or norms and will only act to maximize payoff. And the fact that in this environment of rational interest-maximizing egoists, that a cooperative strategy—tit-for-tat—became the winning strategy is of great significance. His computer programs do not trust anyone, and does not learn to cooperate. It simply follows the simple rules for tit-for-tat and cooperation continues as long as the other players cooperate as well. His research suggests that as long as actors understand that the game is iterated, and that cooperation pays in the long run, then actors will cooperate. There seems to be little need to invoke notions of a norm of reciprocity or trust among players.

Trust and Cooperation: Analysis and Critique of the Rational Actor Model

All that is necessary for cooperation to result in Axelrod's environment is that the initial move of the egoist is to cooperate, and the following moves mirror what the other players did in the previous move. Perhaps one may challenge Axelrod's tit-for-tat strategy by arguing that his model incorporates some degree of trust because the first move of the program is to cooperate. Does this not mean that the program always trusts the other player on the first move? However, I would argue that such a challenge is unfound, for it stretches the meaning of trust. If one returns to Hardin's notion that trust is not strategic, it becomes obvious that the tit-for-tat program is not trusting in the first move, but is simply evaluating what the other player would do by first cooperating. Essentially the program is doing exactly what Hardin writes is not trust: I trust you in order to gain from interacting with you. And after this first move, the program will never trust what the other side does, for it is reactionary. This is the essence of tit-for-tat.

In short, if we assume rationality by all players and assume that the game is an iterated prisoner's dilemma, then trust is not necessary for cooperation to result. Nonetheless, there are limitations to Axelrod's theory even in an environment of interest-maximizing egoists: he does not specify the nature of the good that is being produced from cooperation. In its original form, a prisoner's dilemma is a game where the good that is produced from cooperation—release from prison—is a private good. On the other hand, if the good being produced is a public good, then cooperation may not result even given the payoff structure of Axelrod's game, for it essentially means that each time a tit-for-tat program cooperates, every player in the game will get equal benefits tit-for-tat's cooperation. If that is the case, then the program that cheats the most will end up with the most points, for it will receive the full five points for cheating and also the three points when other programs cooperate. Insofar as most of the goods produced by groups are public goods (private goods do not need cooperation to produce for it can be exchanged on the

market), Axelrod's program may only work for bilateral cooperation. If we are interested in cooperation that occurs in groups of people, then it is necessary to examine more closely the problems of collective action and how people can cooperate too produce public goods.

Like Axelrod, Michael Hechter also works from a rational choice perspective and displays that if groups make side-payments to members in the form of private goods and if groups have adequate monitoring and sanctioning capability, then it is possible for people to cooperate to produce a public good. This is also a scenario where there is cooperation without trust. People will cooperate because they know that it is in the best interest for each member in the group to cooperate given social conditions (such as lack of an alternative group to be a part of) and the structure of the organization itself (mainly the organization's ability to monitor members and apply sanctions to defectors). However, what if there is cooperation among members when there is no good being produced (public or private)? In both Axelrod and Hechter's formulation, cooperation results because there are tangible material benefits to be gained from cooperation. However, there seems to be instances where people cooperate for apparently no benefit.

To solve this problem, it may be necessary after all to invoke the notion of trust. As Tom Tyler argues, there are conceivably two types of trust: instrumental and social. The rational choice perspectives can account for instrumental trust and display that in fact, instrumental "trust" can occur without trust. On the other hand, people care about social status and reputation. Also, people may cooperate with a group because the group is intrinsically a part of their personal identity. For example, according to Craig Calhoun,[1] the student protesters at Tiananmen cooperated to stand up against the state not because they would receive material benefits from

---

[1] Craig Calhoun, "The Problem of Identity in Collecitve Action," in Joan Huber, ed., *Macro-Micro Linkages in Sociology* (Newbury Park: Sage, 1991).

doing so (in fact they faced the real threat of death), but because to not cooperate would be to give up their identity. Their identity came to be founded on the notion that they were a part of a student movement that sought freedom, and to defect would have meant the denial of the self. In this way, non-"rational" concerns motivate people to cooperate in many instances.

The rational choice perspective is indeed useful to display how cooperation may occur in instances where trust is lacking or when actors' only interest is the consumption of a good cooperation can produce. On the other hand, the rational choice perspective is self-limiting in that it can only model such instances. Cooperation oftentimes occurs for no apparent material reason. This is where scholars such as Tom Tyler can offer their insights.

Reo Matsuzaki