

abstract The purpose of rational choice theory is to explain social phenomena by assuming rational choice at the actor's level. To realize this purpose, the theory assumes two mechanisms: choice by actors and the macro–micro–macro transition. After explaining these mechanisms, this article first shows how rational choice theory successfully explains the problem of social order embodied in various fields. It then examines the theory's critics and their related problems. Finally, four new new directions of the theory are proposed: more comprehensive understanding of the interaction between actors and institutions; applications of evolutionary game theory and agent-based modelling; incorporation of the concept of social identity in rational choice theory, and the extension of the coverage of rational choice theory.

keywords agent-based modelling ♦ game theory ♦ micro–macro link ♦ rational choice theory ♦ social identity ♦ social institutions ♦ social order

Overview of theoretical approaches

The purpose of rational choice theory is to explain social phenomena by assuming rational choice at the actor's level (Coleman, 1990; Hechter and Kanazawa, 1997). It has been argued that rational choice theory should explain actors' actions as well. This is discussed later in the article when critiques of rational choice theory are examined and the possibility of incorporating social identity in rational choice theory is explored.

Two mechanisms or processes are analysed to explain how the purpose of rational choice theory is realized: choice by actors and the macro–micro–macro transition. Rational choice theory assumes that an actor chooses an alternative that he/she believes brings about a social outcome that optimizes his/her preference under subjectively conceived constraints. For ease of mathematical modelling, preference is often mapped to a real number, which is called utility or payoff. The above assumption can then be rewritten as follows: an actor chooses an alternative that he/she believes brings about a social outcome that maximizes his/her utility (payoff) under subjectively conceived constraints.

Five critical elements are included in the assumption: constraints, alternatives, social outcomes, utility

and belief. Constraints affect an actor's choices in two ways. First, constraints make some of the possible alternatives impossible. A taboo on sibling marriage, for example, prohibits certain marriage choices, and a ban on smoking prohibits smoking in public places. One cannot buy an airplane if one's budget is only 10 euros. Second, constraints change costs and benefits of alternatives. For example, increasing the price of cigarettes makes the costs of smoking higher. Social approval of hybrid cars increases the psychological benefits of driving such cars.

Note that constraints on the actor are subjectively conceived. As is explained later, the actor has beliefs about the world, and constraints are a part of them. Thus it is subjectively conceived constraints (henceforth, subjective constraints), not objective constraints, that affect the actor's choices of alternatives. Therefore, as stated, the actor chooses an alternative that he/she believes realizes a social outcome that maximizes his/her utility under *subjective* constraints.

Subjective constraints are not independent of objective constraints, however. Social resources such as money, assets, prestige, privilege, authority and power affect the formation of subjective constraints with the help of frames through which the actor views

them. In addition to such social resources, if the actor interacts with other actors, the others' choices become subjective constraints on him/her, and his/her choices become constraints on the others. Game theory, which is a powerful analytical tool for studying interactions between people, properly deals with this situation.

A set of alternatives is available to an actor under subjective constraints. If he/she chooses one of the alternatives, a social outcome is realized. Note that the outcome is not individual but *social*. This means that the outcome is a product of alternatives chosen by more than one actor. Of course, there are cases in which a single actor's choice of an alternative solely determines an outcome. A classic example of these cases is a consumer's choice of goods under his/her income constraint. He/she decides how many apples and oranges he/she should buy at a grocery store with 10 euros. Then his/her decision-making solely determines an outcome: the number of apples and oranges he/she buys. Choices of other consumers and staff of the grocery store do not affect the outcome. Sociologically, however, cases involving more than one actor to produce an outcome are more interesting to study. Thus the adjective *social*, not individual, is added to the noun *outcome*.

The actor then ranks possible social outcomes based on his/her utility. If outcome *A* is preferred to outcome *B*, his/her utility of *A* is larger than that of *B*. If he/she had perfect and complete information on the world and there were no constraints on his/her choice, he/she would choose an alternative that would realize the best social outcome for him/her, that is, an outcome that brings about the highest utility to him/her. However, he/she usually has imperfect, incomplete information about the world, and constraints affect his/her choices.

Information is imperfect and incomplete because the world is uncertain. The uncertainty stems from three mechanisms. First, the actor cannot exactly predict the future of the world. Second, he/she does not precisely know where he/she is now. His/her current situation is a part of a social outcome realized by interactions of involved actors (including him/her). It is often the case that the actor does not know what choices were made by other actors in the past, which contributes to the uncertainty. Third, it is also the case that the actor does not know the utilities of other actors and, therefore, what choices they will make in the future.

Thus the actor has beliefs about the world – about the constraints on him/her, other actors and so on. Then, he/she chooses an alternative based on them. These beliefs are not fixed or static. Rather, the actor updates his/her beliefs using new information about the world. For example, suppose that the actor

thinks that his/her partner in an interaction is uncooperative. If the partner does him/her a favour, he/she would then change his/her belief about the partner's character and behaviour.

The second mechanism necessary for the full-blown analysis by rational choice theory is the macro–micro–macro transition (Coleman, 1990). A macro-sociological explanation explains the occurrence of social phenomenon *Y* by pointing out a precedent social phenomenon *X* that is thought to induce *Y*. Coleman (1990), however, argues that this kind of explanation is unsatisfactory unless the following three processes are clearly uncovered: the macro-to-micro transition, the micro process and the micro-to-macro transition. The macro-to-micro transition is a process in which the precedent social phenomenon *X* creates social conditions for actors such as constraints, beliefs and alternatives. In the micro process actors rationally choose alternatives under the subjective constraints. The choices of the actors then accumulate to produce the social phenomenon *Y* through the micro-to-macro transition.

Take an emergency evacuation of a theatre, for example (Coleman, 1990: Ch. 9). Suppose that a fire breaks out in a theatre and there is only one exit. One possible outcome is congestion at the exit. In this case, the precedent social phenomenon *X* is the fire, and *Y* is the congestion. A causal relation from *X* to *Y* is explained by the following macro–micro–macro transition. First, the fire puts the audience in a new situation and creates two possible actions for them: to take turns at the exit or to rush to the exit (macro-to-micro transition). Second, each person in the theatre decides whether to take turns or to rush (micro process). If all of the people take turns, they can all escape. However, each person has an incentive to rush no matter what action the others choose. If all of the others take turns, rushing to the exit increases the individual's probability of escaping. If all the others rush, the individual actor must also rush in order to not lag behind. Therefore, it is rational for each person to rush to the exit. Third, the actions of rushing to the exit accumulate to cause congestion at the exit (micro-to-macro transition). What is most important in this example is that each person does not rush blindly. Rather, he/she rationally chooses to rush. However, the accumulation of such rational choices produces the congestion.

The social phenomenon *Y* is not confined to congestion in a theatre. It could be the emergence of norms, social structures, inequality, social institutions and so on. Thus rational choice theory has a wide range of *explananda* as long as they are macro-social phenomena. Heckathorn (2001) characterizes rational choice theory as *the interlingua of the social sciences*: the theory provides various fields in soci-

		Player B	
		Cooperation	Defection
Player A	Cooperation	3, 3	1, 4
	Defection	4, 1	2, 2

Figure 1. Payoff structure of the Prisoner's Dilemma

ology with a common language and a universal perspective. Whether the theory succeeds in explaining them, however, is another story, which is discussed later.

Two types of rationality should be explained here. Forward-looking rationality has been assumed so far. This means that actors choose alternatives by predicting the social outcome of their choices. Of course, the actual outcome may be different from the predicted one because of the choices of others, unpredicted events and uncertainty. However, they are assumed to try to choose the alternatives that produce the best outcome for them. Game theory assumes this forward-looking rationality.

The assumption of backward-looking rationality is also used in the literature of rational choice theory. An actor with backward-looking rationality learns from the past. That is, if he/she had previously chosen an alternative the outcome of which was a positive reward, his/her propensity to choose the same alternative again becomes higher. Meanwhile, if the outcome brought about a negative reward, the propensity becomes lower. This assumption is used in evolutionary game theory and agent-based modeling. Furthermore, several types of backward-looking rationality are assumed in social learning models (Fudenberg and Levine, 1998; Young, 1998).

Review of empirical evidence

Critics of rational choice theory often criticize the empirical validity of its assumptions. However, this type of criticism is unproductive from the viewpoint of scientific methodology. It is the hypotheses derived from assumptions, not the assumptions per se, that should be empirically tested. This is the principle of falsifiability proposed by Popper (1959) and elaborated by Lakatos (1970). Some empirical evidence that has survived falsifiability tests is presented below.

One of the fundamental problems in sociology is the problem of social order (Parsons, 1937). How can social order be established among people who have different interests? Rational choice theory has substantively contributed to the solution of the problem. Social order is an abstract concept, though, so this article examines some of the concrete topics embodying the problem of social order: mutual cooperation in the Prisoner's Dilemma game, social dilemma, collective action and the provision of public goods, social movements, social norms and social institutions.

In the Prisoner's Dilemma game, two actors, A and B, who are called 'players' in game theory, have two alternatives (called 'strategies' in game theory) – cooperation and defection. The combination of the strategies chosen by the players produces four possible social outcomes: (cooperation, cooperation), (cooperation, defection), (defection, cooperation) and (defection, defection). Note that the first strategy is of Player A and the second is of Player B. The players receive payoffs depending on the actual social outcome. A conventional payoff structure of the Prisoner's Dilemma game is represented in Figure 1, where larger numbers mean higher payoffs. Player A's payoff is higher if he/she chooses to defect no matter which strategy Player B chooses. This is also the case for Player B. Thus the two players end up in mutual defection to receive the payoff of 2, respectively. However, if both of them choose cooperation, mutual cooperation is realized, and they receive the payoff of 3. Thus mutual cooperation is better for the players than mutual defection is in terms of improving their payoffs. The players cannot realize mutual cooperation, however. They apply forward-looking rationality to their choice of strategies and choose defection individually. Then mutual defection is realized by their *rational* choice. Because of this problem, the game is called the Prisoner's Dilemma.

In real situations, however, mutual cooperation is often observed. Axelrod (1984) presents many exam-

ples of the kind from computer simulation to animal behaviour to trench warfare in the First World War. Thus the theory of the Prisoner's Dilemma game cannot explain the existence of mutual cooperation in the game. The key to solving the puzzle is the iteration of the game. According to Folk Theorem in game theory, mutual cooperation is possible if the game is indefinitely iterated and the players pay enough attention to the future interaction. Axelrod (1984) approaches this issue with evolutionary game theory and computer simulation. The crux of his argument is the strength of the Tit-for-Tat strategy in an iterated Prisoner's Dilemma game. A player with the strategy chooses cooperation at the first round and emulates his/her partner's choice of the previous round. That is, if the partner chose cooperation (defection) at $t-1$, the player chooses cooperation (defection) at t . If the two players adopt the Tit-for-Tat strategy, mutual cooperation is realized at every round. Axelrod's computer simulation and mathematical analysis show that the Tit-for-Tat strategy is evolutionarily stronger than other strategies in that players who adopt this strategy receive higher payoffs on average.

The Prisoner's Dilemma game is played by two players. If more than two players (actors) play a game with a similar structure to the Prisoner's Dilemma game, the situation is called a social dilemma. A social dilemma is defined by two conditions (Dawes, 1980). First, each actor chooses defection no matter how many other actors choose cooperation, because the payoff for choosing defection is higher than that for choosing cooperation. Second, each actor's payoff when all actors choose cooperation is higher than that when all actors choose defection. Following the first condition, all actors choose defection. However, the second condition indicates that this outcome is worse for each actor in terms of his/her payoff than the outcome of universal cooperation. If all actors choose cooperation, their payoff is better than that of universal defection. However, it is rational for each actor to choose defection, which lowers each actor's payoff. Thus this situation is a dilemma. This model explains the mechanism of such social phenomena as the deterioration of natural environments and overfeeding livestock in commons.

A social dilemma is caused by the accumulation of the rational choices of the actors involved. Underprovision of public goods is another type of the accumulation of rational choices that causes 'social irrationality' or unintended consequences. Olson's (1965) theory of collective action captures this mechanism. One puzzle of collective action is why people do not contribute to the provision of public goods even though such public goods would increase their benefits. Suppose, for example, that some resi-

dents are involved in traffic accidents when crossing streets in a community because there is no traffic light. Then installing traffic lights would increase the safety of residents and, eventually, their utility. Thus it is in the residents' interest to form a social movement to put pressure on the local government to install traffic lights. The traffic lights are public goods to residents. However, such social movements by residents are seldom observed even though their interest in traffic lights is great. The installation of traffic lights is therefore not implemented; public goods are not provided. Furthermore, Olson argues, larger groups are more likely to have difficulties in providing public goods.

Why are public goods not provided, especially in a large group? Olson's answer is twofold. First, as the group size increases, the per capita benefit of the provided public goods becomes smaller than the cost per capita for the provision of the public goods. This relation stems from two mechanisms: (1) each actor's share of the benefit of the public goods becomes smaller with the increase in group size and (2) the cost for establishing and managing an organization in pursuit of the public goods increases with the increase in group size. The second reason for non-provision of public goods is the existence of free-riders. An actor would think that the public goods will be provided even though he/she does not contribute to their provision, because the relative effect of his/her contribution is nearly null in a large group. These two reasons lead rational actors to non-contribution to the provision of public goods, which in turn results in the non- or under-provision of such goods.

However, some collective actions succeed in mobilizing people and providing public goods. According to Olson, this can be explained by three elements: small group, selective incentive and coercion. The importance of small groups for the provision of public goods is a natural derivation from the theory outlined above. Selective incentive refers to something provided only to the contributors to the provision of public goods. Examples include T-shirts for participants in a demonstration requiring traffic lights and newsletters only for members of an organization for traffic lights. Coercion can be interpreted as a negative selective incentive.

Social movements are a classic type of the provision of public goods. If a social movement succeeds, everybody gets access to the fruits of the movement even though he/she did not participate in it. However, many people participate in social movements. If Olson's theory is correct, there should be selective incentives. Opp (1989), based on his empirical study of three survey data sets, points out the importance of non-material selective incentives such

as obeying norms internalized in actors. His study has broadened the concept of selective incentive from material to non-material.

The emergence of a new social norm also embodies the problem of social order. An important factor triggering the necessity for a social norm is negative externality of an action. Smoking, for example, decreases non-smokers' utility by damaging their health even though it increases smokers' utility. Thus the negative externality of smoking creates the necessity for a new norm against smoking. However, necessity does not necessarily supply the norm. This is mainly because providing a norm is costly. Promoters of a new norm have to spend time, energy and financial resources to persuade other members in society to accept and follow it. In addition, they may be sanctioned by people who are against it. Thus people who want the norm to be established tend not to promote it; rather, they want other people to promote it. Here we face the free-rider problem again.

Rational choice theory argues that interactions between involved actors are the key to the emergence of the norm. To compensate for the costs of establishing a new norm, people praise the promoters (Ellickson, 2001). The promoters' utility is based on two factors – materialistic and psychological. If a norm against smoking is established in society, the promoters' health will be improved. However, this benefit is also brought to other non-smokers who did not participate in the promotion. This is the mechanism that produces free-riders. The promoters, however, also get psychological benefits because people who support the norm express their gratitude to them. Free-riders cannot enjoy this benefit: it is a selective incentive exclusively provided for the promoters. What is as important as the exclusiveness of this gratitude is that expressing it is not costly to the supporters. Saying 'thank you very much' is nothing to the supporters but something to the promoters.

The emergence of a social institution is another type of the problem of social order. Hechter (1987, 1990) adopts a different approach to Olson's. Hechter focuses on the effect of solidarity on the emergence of a cooperative institution, which, he notes 'enables those who are subject to it to reap a surplus by agreeing on a jointly maximizing strategy that is otherwise unavailable due to the absence or inappropriateness of markets' (Hechter, 1990: 15). For a cooperative institution to emerge in a solidarity group, two conditions must be satisfied. First, members of the solidarity group have demands to consume jointly produced private goods. Second, members have the capacity to prevent free-riding among members and to assure themselves that they intend to cooperate. The first condition seems simi-

lar to the demand for public goods. However, the condition refers to *private* goods. Thus, in Hechter's theory, members are assumed to exclusively consume the goods they provide. As discussed above, demands for a new cooperative institution do not necessarily lead to its emergence; the second condition must also be satisfied. For this condition to be satisfied, Hechter argues, behaviours (production and consumption of the private goods) of members must be visible to other members. This visibility is effective in solving the free-rider problem and assuring members' intentions to cooperate.

Critical assessment of theory and evidence

Critiques of rational choice theory are categorized into three groups. The first criticizes assumptions about preference and utility. The second attacks the empirical validity of the assumption of rationality. The third questions the explanatory power of the theory.

Some scholars in the first group criticize rational choice theory for not asking the origin of preferences of actors. Wildavsky (1992) is representative of this group, pointing out that the Prisoner's Dilemma game becomes a non-dilemma game under certain cultural contexts. He argues that the cultural contexts in which preferences are forged are the focus of cultural theory.

His critique is important in terms of exploring the prior conditions of preferences. However, it is unclear from his argument what social phenomena that have not been explained by rational choice theory are explained by examining the cultural contexts of preference. It is insufficient to merely point out that preference depends on culture. He needs to show that his theory has stronger explanatory power than rational choice theory. Furthermore, his critique does not contradict rational choice theory, but is in fact complementary to it. Even though preference is forged in cultural contexts, the choice process can be studied by rational choice theory once the preference is formed. Thus his arguments concerning the cultural contexts of preference enrich the content of rational choice theory.

Kahneman and Tversky (1979) cast doubt on the assumption of the maximization of expected utility. The assumption is that an actor calculates and compares expected utilities of alternatives using utilities of possible outcomes and probabilities as weights and chooses the alternative that would produce the highest expected utility. Kahneman and Tversky argue that actors use 'prospect' rather than expected utility when they choose alternatives. A prospect is

calculated with 'decision weights' rather than probabilities. However, decision weights are not independent of probabilities. Rather, they are a function of probabilities. This implies that their prospect theory covers a wider range of decision-making processes than the expected utility theory does. Actually, as Kahneman and Tversky (1979: 279) admit, the former generalizes the latter, which, I would argue, enriches rational choice theory. Therefore, the relative explanatory power of the expected utility theory and the prospect theory should decide which should be used for explanation of social phenomena.

The second group of critiques focuses on the empirical validity of the assumption of rationality. Admittedly, the assumption does not perfectly reflect the decision-making of actors. A number of mechanisms – especially psychological ones – must be added to the assumption to perfectly 'explain' the decision-making. However, this line of research would blur the objective of rational choice theory – which is to explain social phenomena. In rational choice theory, actors are assumed to be rational. The theory then tries to explain a target social phenomenon based on the assumption and the analysis of the actors' interactions. If the theory fails to explain it, the assumptions that constitute the theory are to be revised. In some cases, the assumption of rationality needs to be revised. For example, as is discussed later, social identity may need to be incorporated in the theory to enrich the concept of rationality and to enhance its explanatory power. However, criticizing the assumption without referring to the explanatory power of the theory is a fruitless intellectual activity.

Although the first and second groups of critiques of rational choice theory focus on the theory's assumptions, the third group questions the explanatory power of the theory. These critiques are critical to the theory because they argue that rational choice theory does not explain target social phenomena. Green and Shapiro (1994), for example, harshly criticize rational choice theory's accounts of American politics. They focus on four political phenomena – voter turnout, the free-rider problem in political collective action, legislative behaviour and spatial theories of electoral competition. Then they point out the failure of rational choice theories to explain these phenomena.

Although there are some important lessons for empirical tests of rational choice theory in Green and Shapiro (1994), it is impossible to make a point explanation (or prediction) by any theories in social science, let alone rational choice theory. Lovett (2006) points to the importance of comparative statics rather than point explanations/predictions. In comparative statics we calculate how the target variable changes by manipulating a parameter of the

model. Lovett picks up models of voter turnout as an example. The models fail to predict how many voters cast their ballots. (This problem is also called the voter's paradox and will be mentioned again from a difference perspective when the concept of social identity is discussed.) However, Lovett argues that what is more important is the mechanism proposed by the models. Based on the models, voters compare the costs and the benefits of voting, and the voting rate is a simple aggregation of voters who actually cast their ballots. Then a prediction (or a derivation) is that an increase in the cost of voting will decrease turnout. This type of prediction, rather than a prediction of the number of voters casting their ballots, is to be empirically tested.

Future directions

Although there are numerous directions for the future of rational choice theory, four new directions are proposed here. The first direction is a more comprehensive understanding of the interaction between actors and institutions. As mentioned, rational choice theory has contributed to understanding the emergence of institutions. Following Coleman's macro–micro–macro transition scheme, three questions arise. First, how do prior institutions affect actors at the micro level? Second, what is the choice process at the micro level? Third, how do the choices of actors create a new institution? These are core questions in new studies of the interaction between actors and institutions from the perspective of rational choice theory (Nee and Brinton, 1998). The second question was answered above – rational choice by actors. The first and third questions are dealt with below.

In the macro-to-micro transition, prior institutions create social conditions or contexts for actors; they impose constraints on the actors; they provide the actors with new alternatives; they encourage/discourage the actors to have certain beliefs, norms and values; they change the actors' preferences. Actors are not homogeneous, however, in terms of their responsiveness to the effects of prior institutions. Some actors quickly accept a new norm, while others are reluctant to do so. Some actors have more alternatives than others under the same institution. This heterogeneity stems from two elements: the psychological traits and social positions of the actors. More detailed analysis to incorporate these elements is needed.

In the micro-to-macro transition, the emergence of social norms and institutions are to be studied more deeply. Evolutionary game theory and agent-based modelling, which are described in detail below,

have contributed to studies of the spread of particular behavioural patterns such as cooperation and defection in society, and the formation of social structures such as hierarchy and social networks. However, with the exception of book chapters in Hechter et al. (1990), Hechter and Opp (2001) and Nee and Brinton (1998), the emergence of social norms and institutions have not been seriously explored. This is probably because they are conceptually different from behavioural patterns. Suppose that cooperation is prevalent in society. This does not necessarily mean that social norms for cooperation prevail. A social norm includes internalized obligation to conduct a particular behaviour. Prevalent cooperation may have been realized by selfish motivations. If this is the case, social norms for cooperation do not exist in society. Social institutions are also more than mere behavioural patterns. As Coleman (1990) argues, capitalistic behaviours are different from a capitalistic system that consists of the division of labour of corporate actors and institutions such as banks, firms, commodity markets, stock markets and labour markets. Thus it is necessary to clearly conceptualize social norms and institutions and to study their emergence from the perspective of rational choice theory.

A second direction is to advance applications of evolutionary game theory and agent-based modelling. Although it has shown strong analytical power in studying micro-to-macro transitions, conventional game theory finds it difficult to deal with a society with a large population. A player in a large society has to predict choices of many other players in order to choose his/her own strategy, which puts a heavy burden on his/her cognitive capacity. Evolutionary game theory, in contrast, assumes backward-looking rationality rather than the forward-looking rationality that is assumed in conventional game theory. Thus players are assumed to change or not to change their strategies based on their past experience. One more important difference between conventional and evolutionary game theory is that the former focuses on the choices of players (who chose what), while the latter studies the distribution of strategies (how many players chose what) (Skyrms, 1996). The latter can thus be applied to macro-social changes. For example, a change in value in society is a diffusion of a new value taking the place of the old one. An evolutionary game theoretic model can then be applied to the transition process. Players choose one of two strategies: having the old value and having the new value. If having the new value yields a higher payoff than having the old value, the number of players having the new value increases over time, and, eventually, they dominate the society.

This example shows another advantage of ration-

al choice theory. As mentioned earlier, the theory functions as *the interlingua of the social sciences* (Heckathorn, 2001). Although a diffusion model explains the above-mentioned change in value (e.g. Rogers, 1983), the evolutionary game theoretic model analyses the change with a common language used in other fields – rational choice theory. Furthermore, it provides a solid foundation for the diffusion process by explicitly modelling the choice process at the micro level.

Agent-based modelling is an extension of evolutionary game theory (Gilbert, 2008; Macy and Willer, 2002). Conceptually, their differences are not substantive. Both study temporal changes in the distribution of strategies and how society converges (or does not converge) on a particular distribution. In practice, however, agent-based modelling is more flexible than evolutionary game theory. While the latter rigorously analyses models, the former uses computer simulation to produce numerical results. In an agent-based model, some initial properties are assigned to agents. If agents play the Prisoner's Dilemma game in the model, for example, a propensity to cooperate in the game is assigned to an agent. Agents are then paired, play the game and receive payoffs. Agents update their propensities to cooperate based on their payoffs, and they enter the next iteration with the updated propensities.

A distinctive feature of agent-based modelling is its bottom-up approach to social phenomena at the macro level. In a simulation using an agent-based model the researcher manipulates an exogenous variable that represents a prior macro factor in Coleman's scheme. Then a value of the variable becomes a critical initial condition in the model, and agents in the model are set in motion with the condition. Then another macro characteristic, which is the posterior macro variable in Coleman's scheme, emerges from interactions of the agents over iterations. Thus agent-based modelling is a powerful tool by which to study the emergence of the posterior social phenomenon of Coleman's scheme. Importantly, the researcher does not manipulate propensities or behaviours of agents. He/she changes only the values of the exogenous variable, which changes the initial conditions of the agents. Then agents voluntarily behave, and the posterior social phenomenon emerges from their interaction. Thus the self-organization of society is a fruitful objective of agent-based modelling.

A third direction is to incorporate the concept of social identity into rational choice theory (Aguilar and de Francisco, 2009). According to Aguilar and de Francisco (2009), social identity is a social category and a socially distinguishing feature, and there are two approaches to it from rational choice theory:

externalist and internalist rational choice. Externalist rational choice theory does not need the concept of social identity to explain social phenomena, because it deals with the psychological processes of actors as in a black box. Thus it relies heavily on revealed preference. In contrast, internalist rational choice theory focuses on social identity as an important component of the theory. It assumes that an actor conducts a practical syllogism. For example, his/her major premise is 'Democrats vote for the Democratic Party', and, as his/her minor premise, he/she self-identifies as a Democrat. Then he/she votes for the Democratic Party. If actors have social identity of this kind, they would vote for parties they identify themselves with. If rational choice theory successfully incorporates this mechanism of social identity in it, it could solve the voter's paradox. Aguiar and de Francisco (2009) argue that social identity is a set of beliefs about oneself and that, given the identity beliefs, actors make rational choices. Externalist rational choice theory does not include social identity in beliefs held by actors. Internalist rational choice theory, in contrast, explicitly use social identity as beliefs to explain social phenomena that the externalist approach fails to explain. Thus internalist rational choice theory is a promising direction to broaden the horizon of rational choice theory and to enhance its explanatory power.

A fourth direction is to extend the coverage of rational choice theory to explain social phenomena that the theory was not thought to be able to explain. Here I focus on the study of religion. Stark and Bainbridge (1987) open the door to a rational choice theoretic approach to religion and propose a general model of religion based on the theory, and Stark and Finke (2000) study religion assuming a market where religious groups such as churches and sects supply religious 'goods' such as religious services to meet the demand of people. They also assume that the demand for the goods is rather stable, so changes in religious activity are explained by the change in supply.

Turner (2013) casts doubt on the relative stability of the demand. However, Iannaccone (1990) proposes a theory that explains the change in demand at the individual level. He applies theories of household production and human capital (Becker, 1964, 1981) to the study of religious activity. The crux of his argument is that people with religious human capital such as religious knowledge can enjoy religious services and therefore participate in religious activities more frequently and deeply than people without it. The participation, in turn, increases religious human capital. Thus there is an interaction between religious human capital and participation in religious activities. Corcoran (2012) analyses longitudinal

panel data sets in Canada and confirms that her hypotheses derived from the religious human capital theory are generally supported.

Religion is not the only field where rational choice theory shows its powerful explanatory power. Trying to explain social phenomena in other fields by rational choice theory is intellectually challenging but rewarding.

Annotated further reading

- Axelrod R (1984) *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod R (1997) *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton, NJ: Princeton University Press.
- Robert Axelrod (1984) displayed the strength of the Tit-for-Tat strategy by computer simulation. Some critics of his study argued that its strength might depend on the limited number of the strategies used in the simulation. In reply, Axelrod (1997) conducted another simulation of an agent-based model called genetic algorithm and reconfirmed the strength of strategies that share important characteristics with the Tit-for-Tat strategy.
- Boudon R (1982) *The Unintended Consequences of Social Action*. London: Macmillan Press.
- Coleman JS (1990) *Foundations of Social Theory*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Hechter M (ed.) (1983) *The Microfoundations of Macrosociology*. Philadelphia, PA: Temple University Press.
- Raymond Boudon and James Coleman propose the macro–micro–macro transition schemes. Although they use different presentations of the scheme, their substantive contents are the same. Boudon shows some interesting cases in which the rational choices of the actors involved produce unintended consequences. Michael Hechter's edited book collects articles dealing with the transition processes.
- Elster J (1989) *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
- Elster J (2007) *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. New York: Cambridge University Press.
- Jon Elster has published several books on rational choice theory, and these two provide good introductions to the theory. Furthermore, Elster (2007) offers severe, serious criticisms of rational choice theory.
- Hechter M, Opp K-D (eds) (2001) *Social Norms*. New York: Russell Sage Foundation.
- Hechter M, Opp K-D, and Wippler R (eds) (1990) *Social Institutions: Their Emergence, Maintenance and Effects*. Berlin: W de Gruyter.
- Nee V, Brinton MC (eds) (1998) *The New Institutionalism in Sociology*. New York: Russell Sage Foundation.
- These edited books collect excellent articles on the

relationship between individual choices and social entities such as institutions and norms.

Maynard-Smith J (1982) *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Skyrms B (1996) *Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Gilbert N (2008) *Agent-Based Models*. Thousand Oaks, CA: Sage.

Many books have been published on evolutionary game theory and agent-based modelling. John Maynard-Smith opened the door to the study of the society of animals and human beings with evolutionary game theory. Skyrms clearly shows how evolutionary game theory is applied to the social sciences. Gilbert's book is a good introduction to agent-based modelling.

Gibbons R (1992) *Game Theory for Applied Economists*. Princeton, NJ: Princeton University Press.

Gintis H (2009) *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*, 2nd edn. Princeton NJ: Princeton University Press.

As mentioned in the text, game theory is a strong tool for rational choice theory. These books are readable textbooks of game theory. While Gibbons deals with conventional game theory, Gintis provides a clear orientation to evolutionary game theory.

Hechter M, Kanazawa S (1997) Sociological rational choice theory. *Annual Review of Sociology* 23: 191–214.

Heckathorn D (2001) Sociological rational choice. In: Ritzer G, Smart B (eds) *Handbook of Social Theory*. London: Sage.

Scott J (2000) Rational choice theory. In: Browning G, Halcli A, and Webster F (eds) *Understanding Contemporary Society: Theories of the Present*. London: Sage.

Boudon R (2003) Beyond rational choice theory. *Annual Review of Sociology* 29: 1–21.

These articles concisely show the fundamental logic of rational choice theory and its applications in sociology. In addition, Boudon proposes the extension of rational choice theory.

Voss T, Abraham M (2000) Rational choice theory in sociology: A survey. In: Quah SR, Sales A (eds) *The International Handbook of Sociology*. London: Sage.

Rationality and Society. London: Sage.

Although rational choice theorists worldwide have contributed to the progress of the theory, the US, Europe, and Japan are its centres. Voss and Abraham cover excellent works of European rational choice theorists. They also refer to applications of the theory to other social sciences such as economics and political science. *Rationality and Society*, the major journal of rational choice theory, publishes articles of rational choice theorists worldwide.

Green DP, Shapiro I (1994) *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven, CT: Yale University Press.

Friedman J (ed.) (1996) *The Rational Choice Controversy: Economic Models of Politics Reconsidered*. New Haven, CT: Yale University Press.

As discussed in the text, Green and Shapiro made serious criticisms of rational choice theory.

Friedman's edited book collects criticisms of them and their replies to the criticisms. This is a good example of debates on rational choice theory.

References

Aguiar F, de Francisco A (2009) Rational choice, social identity, and beliefs about oneself. *Philosophy of the Social Sciences* 39(4): 547–71.

Axelrod R (1984) *The Evolution of Cooperation*. New York: Basic Books.

Becker GS (1964) *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research.

Becker GS (1981) *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

Coleman JS (1990) *Foundations of Social Theory*. Cambridge, MA: The Belknap Press of Harvard University Press.

Corcoran K (2012) Religious human capital revisited: Testing the effect of religious human capital on religious participation. *Rationality and Society* 24(3): 343–79.

Dawes RM (1980) Social dilemmas. *Annual Review of Psychology* 31: 169–93.

Ellickson RC (2001) The evolution of social norms: A perspective from the legal academy. In: Hechter M, Opp K-D (eds) *Social Norms*. New York: Russell Sage Foundation.

Fudenberg D, Levine DK (1998) *The Theory of Learning in Games*. Cambridge, MA: MIT Press.

Gilbert N (2008) *Agent-Based Models*. Thousand Oaks, CA: Sage.

Green DP, Shapiro I (1994) *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven, CT: Yale University Press.

Hechter M (1987) *Principles of Group Solidarity*. Berkeley: University of California Press.

Hechter M (1990) The emergence of cooperative social institutions. In: Hechter M, Opp K-D, and Wippler R (eds) *Social Institutions: Their Emergence, Maintenance and Effects*. Berlin: W de Gruyter.

Hechter M, Kanazawa S (1997) Sociological rational choice theory. *Annual Review of Sociology* 23: 191–214.

Hechter M, Opp K-D (eds) (2001) *Social Norms*. New York: Russell Sage Foundation.

Hechter M, Opp K-D, and Wippler R (eds) (1990) *Social Institutions: Their Emergence, Maintenance and Effects*. Berlin: W de Gruyter.

Heckathorn D (2001) Sociological rational choice. In: Ritzer G, Smart B (eds) *Handbook of Social Theory*. London: Sage.

Iannaccone LR (1990) Religious practice: A human capital approach. *Journal for the Scientific Study of Religion* 29(3): 297–314.

Kahneman D, Tversky A (1979) Prospect theory: An

- analysis of decision under risk. *Econometrica* 47(2): 263–91.
- Lakatos I (1970) Falsification and the methodology of scientific research programmes. In: Lakatos I, Musgrave A (eds) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.
- Lovett F (2006) Rational choice theory and explanation. *Rationality and Society* 18(2): 237–72.
- Macy MW, Willer R (2002) From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology* 28: 143–66.
- Nee V, Brinton MC (eds) (1998) *The New Institutionalism in Sociology*. New York: Russell Sage Foundation.
- Olson M (1965) *The Logic of Collective Action: Public Goods and the Theory of Group*. Cambridge, MA: Harvard University Press.
- Opp K-D (1989) *The Rationality of Political Protest: A Comparative Analysis of Rational Choice Theory*. Boulder, CO: Westview Press.
- Parsons T (1937) *The Structure of Social Action: A Study in Social Theory with Special Reference to A Group of Recent European Writers*. New York: McGraw Hill.
- Popper KM (1959) *The Logic of Scientific Discovery*. London: Hutchinson.
- Rogers EM (1983) *Diffusion of Innovations*, 3rd edn. New York: Free Press.
- Skyrms B (1996) *Evolution of Social Contract*. Cambridge: Cambridge University Press.
- Stark R, Bainbridge WS (1987) *A Theory of Religion*. New Brunswick, NJ: Rutgers University Press.
- Stark F, Finke R (2000) *Acts of Faith: Explaining the Human Side of Religion*. Berkeley: University of California Press.
- Turner BS (2013) Religion and contemporary sociological theories. *Sociopedia.isa*.
- Wildavsky A (1992) Indispensable framework or just another ideology? *Rationality and Society* 4(1): 8–23.
- Young HP (1998) *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press.

Yoshimichi Sato is Distinguished Professor at Tohoku University, Japan. Sato is the author of 11 books ranging from social change to mathematical sociology and game theory to social stratification. His current research focuses on inequality of income and job security between regular and non-regular workers in Japan, on the emergence of trust between strangers from the perspective of rational choice theory, and on building agent-based models of social networking among individuals. Sato's most recent books include *Japan's New Inequality: Intersection of Employment Reforms and Welfare Arrangements* (2011) and *Social Exclusion: Perspectives from France and Japan* (2012). [email: ysato@sal.tohoku.ac.jp]

résumé La théorie du choix rationnel a pour objectif de rendre compte des phénomènes sociaux en prenant pour hypothèse la rationalité des choix de l'acteur. Elle présuppose alors deux mécanismes: un choix fait par l'acteur et une transition macro-micro-macro. Après avoir explicité ces mécanismes, cet article montre d'abord comment la théorie de l'action rationnelle réussit à expliquer les problèmes relatifs à l'ordre social dans les différents champs sociaux. Il examine ensuite les critiques adressées à cette théorie et leurs limites respectives. Enfin, nous proposons quatre nouvelles pistes de développement pour cette théorie: une compréhension plus poussée du rapport entre individu et institution, applications de la théorie des jeux évolutifs et de la modélisation multi-agent, l'incorporation du concept d'identité sociale dans la théorie du choix rationnel et l'extension de la couverture de la théorie du choix rationnel.

mots-clés identité sociale ♦ institution sociale ♦ lien micro–macro ♦ modèle multi-agent ♦ ordre social ♦ théorie des jeux ♦ théorie du choix rationnel

resumen El objetivo de la elección racional consiste en explicar los fenómenos sociales suponiendo que los actores eligen racionalmente. Para llevar a cabo ese objetivo, la teoría supone que existen dos mecanismos: la elección de los actores y la transición micro-macro. Después de explicar estos mecanismos, este artículo muestra primero cómo la teoría de la elección racional es capaz de explicar las cuestiones relacionadas con el orden social en diferentes ámbitos sociales. A continuación, examina las críticas dirigidas a esta teoría y sus respectivos límites. Por último, proponemos cuatro nuevas vías de desarrollo para esta teoría: una mayor comprensión de la interacción entre actores e instituciones, aplicaciones de la teoría evolutiva de juegos y la modelización basada en agentes, la incorporación del concepto de identidad social en la teoría de la elección racional y la extensión de la cobertura de la teoría de la elección racional.

palabras clave conexión micro–macro ♦ identidad social ♦ instituciones sociales ♦ modelos basados en el agente ♦ orden social ♦ teoría de juegos ♦ teoría de la elección racional