

# Trust and Reciprocity, Interdisciplinary Lessons from Experimental Research

By Elinor Ostrom and James Walker

## SUMMARY



Elinor Ostrom and James Walker wrote *Trust and Reciprocity* on the results of behavioral experiments around trust and reciprocity. Using,

interdisciplinary research, the authors provides diverse perspectives to explain how trust and reciprocity develop. For example, economists, political scientists, psychologists, and sociologists are included.

Using findings from evolutionary psychology, game theory, and laboratory experiments, the authors examine the importance of reciprocal relationships in explaining the origins of trust and trustworthy

behaviour. The common thesis proposes that humans benefit through the development of relationships based on trust. Trust is essential to economic and social transactions that require individuals to behave in cooperative and reciprocal ways.

In five broad sections, this book discusses the motivational drivers for developing trust and reciprocity.

## *Part I: Introduction, Social Dilemmas and Trust*

### **Chapter 1 Introduction**

*Elinor Ostrom and James Walker*

The authors provide an overview of the studies and results of the experiments discussed in the book. This section identifies that social circumstances and context influence trust and reciprocity. The authors suggest that the

conditions that make someone trustworthy should be considered in order to understand trust. In addition, this section discusses that individuals achieve outcomes better than those predicted by game theory models, which are based on selfish motivations. Some people

never trust or cooperate; however, others seem completely selfless. Most people are conditionally trusting, trustworthy, and cooperative based on payoffs of various actions and the intentions of others.

## Chapter 2 Toward a Behavioral Theory Linking Trust, Reciprocity, and Reputation

*Elinor Ostrom*

Ostrom poses the question “How do groups of individuals gain trust? Based on Hobbes’s theory of social dilemmas, Ostrom highlights the value of non-cooperative **game theory** in exploring this question. **Social dilemmas** are defined as situations in which individuals make choices in interdependent circumstances. That is people cooperate in ways to avoid **pareto-inferior outcomes** and move closer to the optimum. Many experiments on trust are based on the **prisoner’s dilemma**, where two people are separated and given the same situation of telling about the other person. If one defects, that person will receive less punishment than the other. If both defect, then they both receive the maximum penalty. If neither defect (or keep silent), they both receive the minimal penalty.



Ostrom indicates that much of this research is based on the **Nash equilibrium**, where each player assumes to know strategies of others and has nothing to gain by changing one’s own strategy. If each player has chosen a strategy and no player can benefit by changing strategies while the other players keep theirs unchanged, then the current set of strategy choices and the corresponding payoffs constitute a Nash equilibrium.

Ostrom identifies the following **six characteristics of social dilemmas** based on non-cooperative game theory:

1. Initial behaviour falls between the social optimal and the sub-game perfect.
2. Cooperation decays slowly with repetition.
3. Communication increases cooperation.
4. The Nash equilibrium is not a good predictor at the individual level.
5. Models based on backward-induction do not predict observed behaviour.
6. Individuals solve second-order dilemmas (who’s going to police the agreement to prevent free riders) to provide rules to improve outcomes.

Ostrom identifies the following behaviours based on the theory of **bounded rationality**:

1. Learn from interactions.
2. Learn to recognize and remember who are trustworthy and untrustworthy.
3. Cooperate with individuals who are expected (from prior experience, visual or verbal cues) to be trustworthy reciprocators.
4. Build reputation of being trustworthy,
5. Punish those who cheated or did not reciprocate.
6. Uses time that extends past the immediate present.

Ostrom suggests that trustworthy behaviour and the ability to solve social dilemmas are due to the interaction of evolved human cognitive capabilities; however, people learn to use their abilities through social experience.

### Chapter 3 Gaming Trust

Russell Hardin

Hardin describes **trustworthiness** as either incentive based (value based) or normatively based (shared interests). Hardin discusses how to measure trust by using the following three situations for understanding cooperation and trust.

*Incentive based trust was highlighted in Mutual Trust and Thick Relationship situations, while normative trust was highlighted in One-way Trust situations.*

**One-way trust** situations are incentive based in which only one person must trust the other to gain from interaction. The truster acts as though they trust the trusted makes acts in their own interest. This leads to cheating by the truster because if they defect and the other cooperates, there are more gains for the truster. The truster is motivated to trust to retain any benefits. Therefore, only when the trusted cooperates, do they both make gains. According to Hardin, people behave trustworthy in one-way situations that are repeated

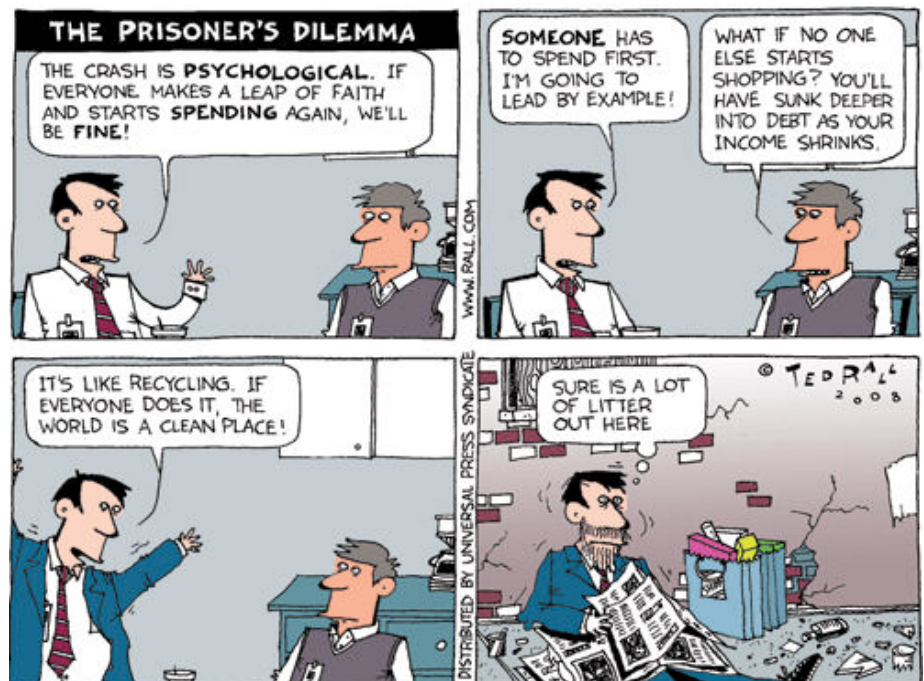
With **Thick Relationships**, interactions occur within an intimate group whereby asocial behavior can be reported to others. For example, in our community everyone is engaged in relationships with each other. People may trust not because they want future interactions with an individual, but are concerned for their own reputation with others. In intimate relationships people trust in order to have the relationship reciprocate and also to preserve reputation.

		Column	
		Cooperate	Defect
Row	Cooperate	20,20	-25,40
	Defect	40,-25	0,0

A **Mutual trust** situation occurs with the prisoner's dilemma, whereby both participants move at the same time. Therefore, both people have to trust the other in order to benefit from the interactions. Hardin suggests that in mutual situations people are trustworthy because they do

not want the other party to pull out of a beneficial relationship. Hardin suggests that simultaneous cooperation involves incentive based trust. Mutual exchange can be represented as a Prisoner's Dilemma game; both participants get the most benefit when they both cooperate. In mutual situations people are trustworthy because people do not want the other party to pull out of a beneficial relationship.

because people care about future interactions. For example, in a game where both are required to invest money, the truster must cooperate for the trusted to receive gains, but the trusted need not cooperate back, therefore the relationship is one-way.



## **Part II: Biological Foundations of Trust and Reciprocity**

### **Chapter 4 Biological Foundations of Reciprocity**

*Robert Kurzban*

Kurzban uses Robert Triver's idea that **altruism** exists as a mechanism of cooperation, in contrast to natural selection theories. Altruism is defined as behaviours that benefit another organism, not closely related, while being apparently detrimental to the organism performing the behaviour. Kurzban suggests that natural selection can lead to phenotypic features designed to deliver benefits to others through **kin selection** and **reciprocal altruism**. Kin-selection strategies in evolution favour the reproductive success of an organism's relatives, even at a cost to the organism's own survival. Kurzban suggests that the process of natural selection obeys game theory in that genes persist by rate of replication. The resulting cognitive systems the genes build do not have to be rational, but are better than others.

The pure evolution of gene sequencing does not address the fact that there is apparent altruism in nature. Therefore, Kurzban identifies **reciprocal altruism** in which beneficial outcomes occur in repeated social dilemmas based on conditional strategies. For example, **tit-for-tat** is a conditional strategy where one agent will first cooperate, then

subsequently replicate an opponent's previous action. If the opponent previously was cooperative, the agent is cooperative. If not, the agent is not.

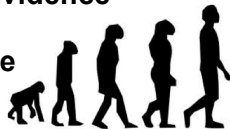
#### **Reciprocal altruism involves the following conditions:**

- 1. The environment must be one in which there are benefits to be conferred.**
- 2. Organisms must have repeated interactions with one another.**
- 3. Organisms must have ability to distinguish those who they will benefit from others.**
- 4. Organisms must have sufficient behavioural flexibility to interact with others.**

The process of natural selection generated cognitive abilities and decision rules that lead to benefits of some over others. Kurzban suggests that these biological mechanisms help identify the payoffs in social dilemmas. For example, our memories are able to keep track of good partners who play by the rules and bad partners who cheat. Thus reciprocal altruism is an evolutionary pathway by which benefits may be derived from a relationship and trust can build upon.

### **Chapter 5 The Chimpanzee's Service Economy: Evidence for Cognition-Based Reciprocal Exchange**

*Frans B.M. de Waal*



Frans de Waal describes and expands on the biological foundations of trust by exploring the relationship between **evolution, cognition, and behavior**. De Waal studied chimpanzees grooming and food sharing patterns. He observed that chimpanzees demonstrate reciprocity in both cases. For

example, Chimp A's sharing of food or grooming of Chimp B correlated with Chimp B's sharing with Chimp A. A chimp will scratch another's back if their back is scratched. Thus, Chimpanzees use biological mechanisms such as tit-for-tat to sustain relationships.

A **theory of evolution of social behaviour** is based on the link between humans and chimpanzees, as they are most genetically similar. Although chimpanzees are less cognitively

evolved than humans, reciprocal and trusting behaviour can be identified among them. This suggests there are evolutionary mechanisms involved in the development human behaviour.

***Demonstrated as primitive reciprocity in chimps, humans have evolved the capacity to remember the actions of others, which is manifested in trust.***

## **Part III: The Links Between Evolution, Cognition, and Behaviour**

### **Chapter 6 A cognitive Theory of Reciprocal Exchange**

Kevin A. McCabe

McCabe indicates that reciprocity behaviours are generated by cognitive strategies implemented in the evolved neural circuitry of the nervous system. He studied **reciprocity at four levels**.

- 1. Behavioural level:** Looking at the messages and rules of the messages.
- 2. Strategic level:** Observing variations in reciprocal behaviour, decision making, utility, and rationality.
- 3. Evolutionary level:** Looking at those strategies that have survived throughout evolution.
- 4. Neural level:** How the nervous system encodes strategies.

At the behavioural level, **positive/negative reciprocity** are observed in one-shot games. Positive reciprocity are reciprocal tendencies or inclinations to cooperate. Negative reciprocity includes aspects of trying to retaliate and cause harm. He also identifies a mechanism known as **goodwill accounting**, whereby people keep track

### **Chapter 7 Conflict, Interpersonal Assessment, and the Evolution of Cooperation: Simulation Results**

James Hanley, John Orbell, and Tomonori Morikawa

The authors discuss game theory studies on the role of conflict in assuring cooperation and trust in social dilemmas.

**Trust is a subjective expectation that a person will act in a cooperative manner.** These expectations are a function of the information available about others and their ability to process it. The

research shows that if people have the option of conflict with others, there will be more pressure to try to see if others can be trusted and therefore cooperate with them. The **Hawk-Dove game** is an alternative to the prisoner's dilemma to investigate trust and conflict. Players choose to pursue a good and risk conflict or adopt a safer strategy. The "hawk" generally initiates conflict and a "dove" will avoid conflict. If two hawks meet, there will always be a conflict. Winners receive the benefit and losers face the cost of the conflict. Doves flee, and are never involved in a conflict.

of which partners can be relied on to establish trust. This is a robust strategy for positive reciprocity situations and is necessary for **reciprocal altruism** to occur. McCabe also notices that reciprocity depends on **social distance** or the degree to which people feel connected or empathize with others. For example, Person 1 is given money and can keep all of the money, give some away, or give all away to Person 2. The more social distance, the less money people give to others. In evolutionary psychology, this concept is reinforced, as primates would groom each other or share food more if there was a mutual benefit. However, especially with food sharing, people reciprocate more with closer social distance.

McCabe links reciprocal behavior in games to the part of the brain that has evolved along with social skills. Using trust games, he illustrates the cognitive and neurological aspects of decision-making by imaging people's brains by magnetic resonance imaging (MRI) scanners. MRI images demonstrate differences in the density of blood flowing to different regions of the brain. When people are in situations involving **social reasoning**, the **prefrontal cortex** is shown to be intense in the imaging. Other regions of the brain are active with reciprocal behavior in games or with human social behaviour.



There is no cost to be a dove, except not receiving a payoff. As the ratio of cost to benefit increases, the population of hawks will decrease. This research demonstrates that **cooperative dispositions will be more likely to evolve with conflict.**

## **Part IV: Experimental Evidence**

### **Chapter 8 Experimental Studies of Cooperation, Trust, and Social Exchange**

*Karen S. Cook and Robin M. Cooper*

Cook and Cooper suggest that trust facilitates cooperation and social order. The authors discuss social psychological experiments on the foundations of trusting relationships. The authors reference Morton Deutsch research with the prisoner's dilemma. Deutsch (1960) suggests that there are **three types of players** who are motivated by their conditions or contexts. They are the **confrontational, individualistic, and cooperative**. That is, those who are concerned about other people's welfare are most inclined to cooperate and make a "trusting choice". Other studies such as Rotter (1967, 1971) explore the extent of trust as a personality factor that can be used as a predictor of social behaviour. For example, Rotter found that high trusters are trustworthier, have difficulty in lying and more likely to contribute to upholding public goods. Yamagishi designed a variant of the prisoner's dilemma game in different cultural contexts to allow people to adjust the amount that is at stake. This game helped to highlight what erodes or builds trust in different people.

The authors suggest that there are **three core elements to cooperation**:

- 1. Motivations**
- 2. Incentive Structures**
- 3. Social Context**

**Motivations:** Social Psychologists discuss motivations for cooperation. Egoists maximize their own payoffs, cooperators minimize the difference in payoffs, and altruists maximize the other's payoffs. Egoists are also the least likely to trust others. Motivations also include the expected behaviour of others. The authors conclude that high trusters are more accurate at determining other's behaviour.

Another experiment using computer simulation showed that higher levels of trust benefitted cooperators when there was high opportunity cost; however, the amount of cooperation depended on their level of trust. Also, further studies showed that high trusters are less fearful of others (cheating potential) than low trusters.

**Incentive Structures:** These are connected to strategies used in situations. Some examples include tit-for-tat and hostage posting. **Hostage posting** occurs when subjects offer part of their payoffs as incentive for their partners to reciprocate trust. This increases the payoff for the other and the cost if they defect. For example, if one person accepts a watch for the cost of a loan. If the person does not pay back the loan, the other person gets the watch. **Thus, hostage posting is an incentive mechanism to reach the paereto-optimal scenario.** Generally, those who post a hostage, trust their partners. Also, trust decreases with increased risk. As temptation to defect increases, there is less likely that trust is honoured.

**Social Context:** Group identity, group size, communication, time pressure, and culture influence trust. For example, Yamagishi (1988) suggests that North American culture is more individualistic, versus Japanese culture, which is more collective.

There were several experiments on the relationship between **trust and social exchange**. For example, more people trust with uncertainty and with more frequent game partners (relational commitment). Uncertainty allows participants to build trust because they need to exchange resources/ideas in order to benefit. Finally, **trust can be negotiated** as demonstrated by the fact that cooperation rates were higher with structured games. Even with no assurance of return, players trusted more when the game disallowed free riding.

## Chapter 9 The Human Face of Game Theory: Trust and Reciprocity in Sequential Games

Catherine C. Eckel and Rick K. Wilson

The authors suggest that **humans share a capacity to read one another's intentions through a set of cues such as facial expressions, body language, and tone of voice**. This ability has evolved with the mechanisms for displaying intentions. The authors discuss how game theory is useful for reading intentions. For example, in bargaining situations most people choose strategies that appear to be cooperative. Players must also draw inferences about the trustworthiness of the other. One study on autistic people indicated that people often show signals that betray intentions. Autistic people cannot read facial expressions and therefore cannot develop a "theory of mind" or be able to put themselves in the others shoes.

**Theory of mind** involves the ability to detect others purpose, uses eye direction to recognize the focus of intention and the ability to connect two individuals to a situation or object. For example, one person can follow the gaze of another to a bottle of beer. They then connect eyes and there is a realization that one sees the other looking at the bottle. The intention of the onlooker is revealed and the gazer shares the beer.

**Emotion** can add credibility to the signal of intention. For example, anger cannot be easily enticed, but emphasizes the intention to defect.

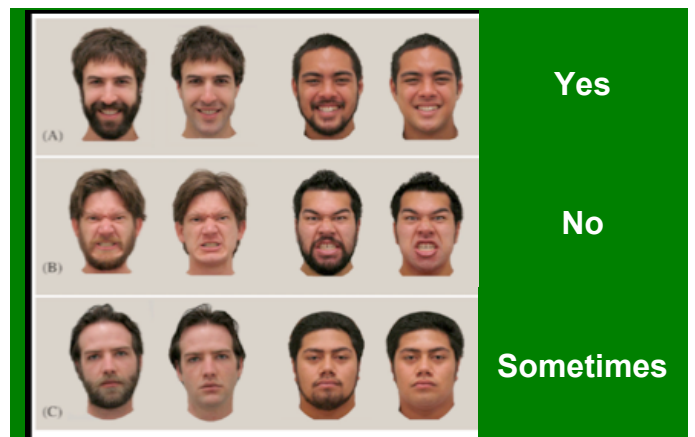
**Social signals**, depending on cultural context, help reveal intentions as well. For example women may be more cooperative than men.

**Facial expressions** can also emphasize intentions.

The authors discuss two experiments that tested whether people can discriminate among partners by facial cues. Smiles are the most recognized expression of happiness. The studies show that people who look happy tend to be perceived as trustworthier or less trustworthy. For example, the results show that a smiling male was viewed as cooperative; however, a neutral female was viewed as more genuine than the smiling one. In addition, smiling in general helps build trust with strangers.

Finally, females and males respond differently to smiles, in that females were less likely to trust other smiley females than males did. Thus, the studies show that smiles do not always indicate trustworthiness; however, smiles do invite trust. Therefore people have the capacity to read intentions behind cues, but are conditional. Thus, the ability to read intentions helps in social dilemmas.

### Perception of trust by Facial Expressions



**Trust is dependent on the complex relationships between incentives and individual characteristics. Social contexts can promote or diminish trust.**

## Chapter 10 Strategic Analysis in Games: What Information Do Players Use?

Kevin A. McCabe and Vernon L. Smith



McCabe and Smith discuss **mental mechanisms** that people use in social exchanges for whether they cooperate or defect. Through repeated interactions with different partners, people adapt through learning to play non-cooperative equilibriums. According to standard game theory, cooperation increases with the probability that subjects will be matched with the same partner; however, it does not account for the pattern of defection and punishment across experiments.

The authors critique game theory in that the concept of cooperation pays and defection does not pay is too simplistic for the variety of behaviour that can be observed. Therefore, the authors suggest using **two-person bargaining experiments** to identify players with variable degrees in likelihood to defect or cooperate with others offering positive reciprocity.

The results show that people use a form of **goodwill accounting** whereby people track trading partners to predict reciprocal behaviour in receiving mutual gains or intentionality of defecting. By observing both game types, the authors suggest **that people prefer to exchange like things such as monetary for monetary or favours for**

**favours**. Also, goodwill accounting allows participants to have triadic knowledge. For example, a person may know that the other knows that he is hungry and then offers food. However, there is an expectation for reciprocity.

**Game 1:** In this game there are high payoffs and punishment. Each player alternates moves, beginning with player 1, until a payoff is reached. Certain actions may result in different random payoffs such as a 60,40 or 40,40 split or a perfect 50,50 split. Each player can punish the other if they defect. Players often overcome the risk of defection by displaying goodwill accounting in order to maximize the relationship through cooperation. Game 1 is a normal form of games that use typical matrices and used in identifying strategies that result in the optimal gains.

**Game 2:** Unknown to the players, the payoffs are interchanged so that Player 1 can only receive 60,30. The game then relies on trust because player one can defect unless there is positive reciprocity. These extensive games involve the sequencing of players' possible moves, their choices at every decision point, the information each player has about the other player's moves when he makes a decision, and his payoffs for all possible game outcomes.

## Chapter 11 Trust in Children

William T. Harbaugh, Kate Krause,  
Steven G. Liday Jr., and Lise  
Vesterlund

The authors assess the how trust and reciprocity are learned behaviors. Their research involved children of different ages to play trust games. Children from grades 3, 6, 9, and 12 play extensive form trust games with an anonymous member from each of the other grades. The studies use **one-**

**shot dictator games** in which one player is asked to decide how much of the pie to give to another anonymous player. In **public good games**, one would sacrifice payoffs for the good of others.

The authors find that children trusted older people more than within their own cohort, suggesting that status matters. Also, children are generally less trusting than adults. It is possible that older children have

experienced more positive reciprocity and are more trusting. However, children ages 6-12 contributed most to the public good. The authors suggest that adults have learned strategies to assess payoffs more appropriately and have also learned to free ride. In addition, older children realized that their own trust could be violated or that they could exploit others, suggesting that trust develops with age.





## Chapter 12 Trust in Two-Person Games: Game Structures and Linkages

T.K. Ahn, Elinor Ostrom, David Schmidt, and James Walker

The authors discuss

the effects of changes in payoff structure of social dilemmas and in the length of interactions using both one-shot and repeated prisoner's dilemma experiments. For example, if Player 1 can give \$5 to Player 2 and Player 2 gives \$5 back, they both receive \$10. If noone gives to the other, then they keep \$5 but receive nothing more. If Player 1 defects and Player 2 gives money, Player 1 gets \$15 and Player 2 gets \$0. **Players in one-shot games tend to be precautionary and are more likely to defect. Players in repeated games are more likely to cooperate.** This suggests that that there

## Chapter 13 Cross-Societal Experimentation on Trust: A Comparison of the United States and Japan

Toshio Yamagishi

信

Yamagishi summarizes his research on trust differences between Japan and the United

States. It is generally believed that Japanese society, in particular in the business sector involves high levels of trust versus the individualism of the United States. However, Yamagishi demonstrates that **Americans are just as trusting.** People use **assurance** when they respond in ways that benefit others so they can also benefit. In contrast, people

behave in **trusting ways** when they act to benefit others because the care for them. Japanese do not trust more, but they can construct assurance better. Market experiments simulating transactions of rice and rubber showed that the **nationality of people did not contribute to levels of commitment.** Rather, **social uncertainty** and **opportunity cost** were more influential. For example, for both nationalities, the less people knew about the other players, the less trust would result. People generally trusted more often when they expected others to cooperate than defect. Regardless of nationality, if there was more at

are several factors (box) that influence trust. **Overall, those who receive direct information about the other player are more likely to cooperate especially if the other players are more aligned with their own values.** It was also found that historical trusting behaviour was positively associated with current trust. Finally, cooperation leads to decreasing greed and fear with payoff differences only mattering when players can judge the trustworthiness of others based on available information.

**Precuniary benefits (utility):** They are the relations' among precautionary payoffs. Even in a one-shot game most people defect no matter the level of fear, greed, or cooperators gain. Greed is the difference between the payoff for defecting and the payoff for cooperation. Fear is the difference between the payoff of defecting on a defector and being exploited by a defector.

**Types of players:** Due to different values, players have different perception on payoff gains. For example, some players believe gains are more than monetary. Cooperation can be a gain for some.

**Information available** about the player types: For example "nice" players cooperated more when paired with other "nice players" than with "stingy" players.

**Linkages among players:** In repeated games, if players believe that the persons whom they are matched could be influenced by current decisions, such linkages could affect behavior.

stake, there was higher chance of defecting. Therefore, it was more important in situations of high uncertainty to build trusting relationships. However, **Americans have a higher level of general trust.** Therefore, their expectations that others would cooperate are based on a general belief that people do. The Japanese would expect cooperation from the partner within the relationships in which mutual monitoring and control are possible (supporting the collectivism theory). According to the author, this type of trusting is not useful in one-shot games, but can be observed in repeated games. Finally, the Japanese require more of an illusion of control (more information).

## **Part V: Conclusions**

### **Chapter 14 The Transformation of a Sceptic: What Nonexperimentalists Can Learn from Experimentalists**

*Margaret Levi*

Levi suggests that laboratory experiments are flawed in the sense that there is no external validity. However, she recognizes that experimental research, although unorthodox and not easily understood, adds value in tackling complex topics. Levi indicates that game theory provides a normative account of rationality and provides opportunity for backward deduction. Levi's main critique of game theory is that games only provide a null hypothesis in that they use prescribed conditions and therefore under-predict reciprocity. Levi also highlights that most of the research is about trust and not reciprocity; however, trust is important for reciprocity to occur. Finally, she suggests that social context is important in understanding any behaviour, even in experiments.

### **A Brief Critique**

I agree with Margaret Levi that the conditions in experiments using game theory do not represent every possible situation in life. However, this book provides a robust overview of the underpinnings of trust and reciprocity derived from behavioral experiments. By applying game theory to broad interdisciplinary research such as psychology, biology, and economics, the authors illustrate that human's benefit through the development of relationships based on trust. In addition, Ostrom and Walker have shown that trust is essential to reciprocity. The results of the research did not agree with the expectations in game theory, in that there were a great variety of behaviours. For example, some people cooperate, some people do not cooperate or trust, and some are altruistic. People have evolved both biologically and socially to adapt to



### **Chapter 15 Conclusion**

*James Walker and Elinor Ostrom*

Walker and Ostrom summarize some overall themes of the book. For example, experiments using chimpanzees as subjects demonstrate that reciprocity is

based on exchange mechanisms which direct social behaviour that require the services from others (usually same partner). This supports the evolutionary origin of the capacity to learn to trust. The authors suggest that people use cues outside the game theory experiments to assess trustworthiness, such as historical behaviour and physical expressions. Therefore, trust is conditional. Also, social distance is a key variable in explaining deviations based on self-interested payoffs. Ostrom and Walker conclude with the notion that the interdisciplinary experimental research in biology, technology, economics, philosophy, and psychology have provided an **adaptive toolbox** for understanding trust and reciprocity with different social norms, institutions, and other cultures.

situations that result in benefits; however, there is still heterogeneity in trusting behaviour. Therefore, the book highlights that social and cultural contexts are important in trust and reciprocity. Thus, the authors identify that trust is the main challenge for cooperation. Further, the authors emphasize that reciprocity requires one to move beyond selfish behaviour. This research could be expanded to include the importance of trust and reciprocity in leadership.

#### **Discussion questions**

1. How can this research in trust and reciprocity be applied to leadership?
2. Does cooperation always require trust and how authentic is it?

### **Authors**



Elinor Ostrom is a professor of political science and co-director of the Workshop in Political Theory and Policy Analysis, Indiana University, Bloomington. She was awarded the Nobel Prize in economics in 2009.

James Walker is a professor of economics, former chair of the Department of Economics at Indiana University and former co-director of the Vincent and Elinor Ostrom Workshop in Political Theory and Policy Analysis, Indiana University Bloomington.

