

JÖRG RIESKAMP and PETER M. TODD

## THE EVOLUTION OF COOPERATIVE STRATEGIES FOR ASYMMETRIC SOCIAL INTERACTIONS

**ABSTRACT.** How can cooperation be achieved between self-interested individuals in commonly-occurring asymmetric interactions where agents have different positions? Should agents use the same strategies that are appropriate for symmetric social situations? We explore these questions through the asymmetric interaction captured in the indefinitely repeated investment game (IG). In every period of this game, the first player decides how much of an endowment he wants to invest, then this amount is tripled and passed to the second player, who finally decides how much of the tripled investment she wants to return to the first player. The results of three evolutionary studies demonstrate that the best-performing strategies for this asymmetric game differ from those for a similar but symmetric game, the indefinitely repeated Prisoner's dilemma game. The strategies that enable cooperation for the asymmetric IG react more sensitively to exploitation, meaning that cooperation can more easily break down. Furthermore, once cooperation has stopped, it is much more difficult to reestablish than in symmetric situations. Based on these results, the presence of asymmetry in an interaction appears to be an important factor affecting adaptive behavior in these common social situations.

**KEY WORDS:** bargaining, evolutionary stable strategies, finite state automata, investment game, repeated games.

### 1. INTRODUCTION

The study of cooperation between self-interested individuals has been dominated by a focus on symmetric games. But social relationships are often imbalanced. Consider for instance the relationship between an employee and employer or a patient and a physician. In the first case, an employee might have frequently worked overtime trusting that his effort will be reciprocated, but the employer may or may not ultimately reward this extra effort. Individuals must frequently deal with others in

positions at a different level of power, making their interactions asymmetric. Whereas the two individuals in a symmetric interaction face the same decision alternatives, in an asymmetric interaction the decision alternatives differ for the two parties. Do these commonplace asymmetric interactions support or inhibit the emergence of cooperation?

The main goal of the present study is to explore whether decision strategies exist that can reliably lead to ongoing cooperation in an asymmetric interaction, and see whether such strategies differ from those that are appropriate for cooperation in a symmetric interaction. To achieve our goal, we first narrow our attention to a particular simple form of asymmetric social exchange, captured in the *investment game* (IG—Berg et al., 1995). This game, sometimes also labeled the *trust game*, has so far been studied experimentally, but here we take a simulation approach. This approach begins with the assumption that human cognition is adapted to the problems people face, implying in particular that people possess or learn strategies that produce beneficial outcomes in social interactions. Consequently, decision strategies that compete well against other strategies are more likely to be used. We will study the way in which strategies compete against each other by means of simulations of evolutionary processes acting on populations of agents. These simulations proceed as a sequence of generations in which agents equipped with specific strategies interact with one another. Agents with better-performing strategies—that is, those that get higher payoffs compared to others—are more likely to be present in subsequent generations. During the evolutionary process the strategies can also be modified (e.g., by mutation) and thereby potentially improved, and ultimately the simulation often converges to a set of high-payoff strategies.

The results of such evolutionary simulations can be employed to help explain how cooperation can develop and to illustrate strategies that guide people's decisions. We describe these strategies in the form of finite automata (e.g. Abreu and Rubinstein, 1988; Aumann, 1981; Binmore and Samuelson, 1992; Nelson, 1975; Suppes, 1969), which are models that change their state

and their corresponding output depending on the inputs encountered. We compare the strategies we find for the repeated IG with those that have been investigated for the repeated symmetric *Prisoner's dilemma game* (PDG—Luce and Raiffa, 1957), usually taken as the standard framework within which to study the emergence and breakdown of cooperation in social interactions. Given that asymmetric situations are common, as we argue, the IG provides an extension of the studies previously carried out on the PDG.

This article is organized as follows. We first introduce the IG as a simple form of an asymmetric interaction, and then use three game-theoretical concepts to evaluate the performance of a small set of strategies for the IG. Through this evaluation the weaknesses of the game-theoretical concepts and the need for our evolutionary approach, manifested in the three studies that follow, become clear. In the first evolutionary study the differences between the IG and the PDG are examined. The second study explores the effect of strategy complexity on performance. The third study investigates the consequences of the sequential decision process in the IG. Finally, we discuss the results of all three studies to address our guiding question of how particular strategies can explain cooperation in an asymmetric interaction.

## 2. EVALUATING THE PERFORMANCE OF REPEATED GAME STRATEGIES

What are strategies that enable cooperation—trust and reciprocity—between unrelated individuals? As a first step to understand the mechanisms that enable cooperation, we will evaluate strategies analytically, illustrating their strengths and weaknesses. For this purpose, first an asymmetric social interaction is defined.

### 2.1. *The investment game representing an asymmetric social interaction*

In asymmetric games, player roles are not interchangeable as they are in symmetric games—that is, the two players face

different decision options with different resulting payoffs. The IG represents an asymmetric interaction between two individuals. In this two-person sequential bargaining game, both players receive an endowment. First, player A can invest any amount of the endowment, which is then tripled, producing some surplus before it is delivered to player B. Next, player B decides how much of the tripled amount she wishes to return to player A (Figure 1).

Assume each player starts with an endowment of 10. Player A could invest 5 of his 10 endowment, which is then multiplied to 15 and sent to player B. Player B could then return 5, so that player A ends up with the five he kept and the additional five player B returned, and player B receives the 10 of the trebled investment she kept and her own endowment of 10. If player A makes any investment, this leads to a total payoff for both players greater than the sum of both endowments, producing a *surplus*. If player A invests the whole endowment the maximum possible surplus is produced, defining an *efficient* outcome, whereas any investment lower than 100% leads to an *inefficient* outcome. The game-theoretic prediction (the subgame-perfect equilibrium, see Fudenberg and Tirole, 1991) for the IG is straightforward: to maximize her monetary payoff, player B will return nothing to player A. Player A can anticipate this, so he will not make an investment, which leads to the most *inefficient outcome* with no surplus.

However, experimental results deviate from this equilibrium prediction. In a study by Berg et al. (1995), participants in the position of player A sent on average \$5.2 of their endowment of \$10 to their counterpart, whereas participants in the position of player B returned on average \$4.7 of the trebled investment. Only two of the 32 player A participants in the experiment sent nothing to their counterpart. This result is in line with other experimental studies of asymmetric games showing that people often trust each other and reciprocate trust with fair decisions (Fehr et al., 1997; Güth et al., 1997; Van Huyck et al., 1995).

How can this observed cooperation be explained when assuming individuals are primarily self-interested? One common

explanation is based on the idea that *repeated* interactions enable cooperation by creating the opportunity to reciprocate cooperative behavior and punish uncooperative behavior. Trivers (1971) was one of the first who emphasized the role of repetition, which has attracted most of the research that tries to explain cooperation between unrelated, self-interested individuals (Axelrod, 1984; Schelling, 1960; Sugden, 1986). Beside repetition other factors that promote cooperation between self-interested individuals have been studied, for instance the role of reputation (Alexander, 1987), agents' interaction depending on spatial localization (Hoffman, 1999; Nowak and May, 1992; Aktipis, 2004), or the self-commitment toward cooperation detectable by others (Güth and Kliemt, 2000).

2.2. *Repeated games*

Indeed, repetition can support cooperation in the symmetric indefinitely repeated (iterated) Prisoner's dilemma game (IPD). The two players of the unrepeated, "one-shot" PDG can either cooperate or defect (see Figure 2). For both players, defecting always leads to a greater payoff regardless of whether the opponent defects or cooperates; thus, defecting is a dominant strategy. However, each player receives a greater payoff when both players cooperate than when they both defect, thus the individual payoff-maximizing strategies (mutual defection) lead to an inefficient outcome.

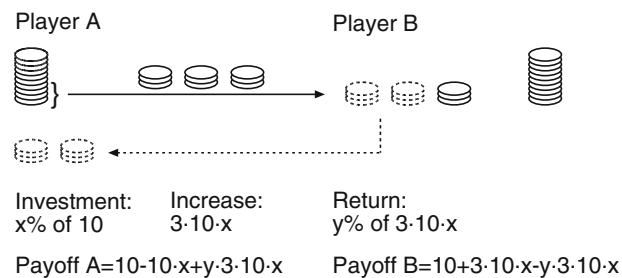


Figure 1. The investment game. The payoff of player A is defined as his endowment minus his investment plus player B's return. Player B's payoff is defined as her endowment plus the trebled investment of player A minus player B's return.

Player A	Player B	
	Cooperate	Defect
Cooperate	20, 20	0, 40*
Defect	40*, 0	10*, 10*

Figure 2. The Prisoner's dilemma game. The first value in the cells represents player A's payoff and the second value player B's payoff. Because the game is symmetric, the roles of players A and B are interchangeable. The asterisks indicate the best reply strategies for each player. The cell with two asterisks indicates the Nash equilibrium.

The PDG differs from the IG in its symmetric payoff structure, simultaneous decisions, and dichotomous decision alternatives (cooperate/defect versus a numerical amount to transfer). However, both games represent a social dilemma, in that the individuals' rational decisions lead to a socially undesired (and inefficient) outcome—the absence of cooperation. Due to this similarity between the two games, we begin our exploration of cooperating strategies for the repeated IG by considering strategies that have been studied for the IPD.

In the IPD the game is repeated for an unknown number of periods, so that cooperating can become payoff-maximizing because the other player may reciprocate cooperation in the hopes of maintaining the benefits of mutual cooperation in future periods. Axelrod (1984) demonstrated that a simple strategy called *Tit for Tat* (TFT) can lead to cooperation in the IPD, as it cooperates as long as the opponent cooperates and defects only if the opponent defected in the previous period. Astonishingly, this simple strategy outperformed many other more sophisticated strategies in tournament competitions. Another simple strategy, *Grim*, cooperates in the first period and continues to cooperate unless the other player defects, in which case it defects in all following periods regardless of the opponent's decisions (Binmore, 1994), which makes Grim less cooperative compared to TFT. The main advantage of Grim is that it provides little opportunity for exploitation as it can only be exploited once, which can make it competitive in evolutionary simulations (Linstler, 1992).

However, both Grim and TFT are vulnerable to mistakes. For instance, if two players apply TFT and one player defects accidentally, this leads to a long period of alternating defection,

until one player makes a second mistake and mutual cooperation returns (Nowak and Sigmund, 1993). The strategy *Pavlov* (or *win-stay/lose-shift*) out-competes TFT if small mistakes occur. Pavlov cooperates in the first period and then cooperates if both players made the same decision in the previous period, otherwise it defects (Kraines and Kraines, 1993, 1995). If one of two players cooperating by applying *Pavlov* defects by mistake, both players will defect in the following period, but will proceed to cooperate again in the period after that. Other strategies less vulnerable to mistakes are *Generous TFT* (Nowak and Sigmund, 1994) or *Contrite TFT* (Boyd, 1989; Sugden, 1986). Non-cooperative strategies also frequently arise in evolutionary simulations, as for instance the strategy *always-defect*, which defects regardless of the opponent's decision. It outperforms TFT when playing against it in a direct competition. Starting from a heterogeneous population of randomly created strategies, *always-defect* is often the strategy that prevails in the population at the beginning of an evolutionary process before other strategies evolve (Boyd, 1989; Hoffmann, 2001; Marinoff, 1990; Sober, 1992; Young and Foster, 1991).

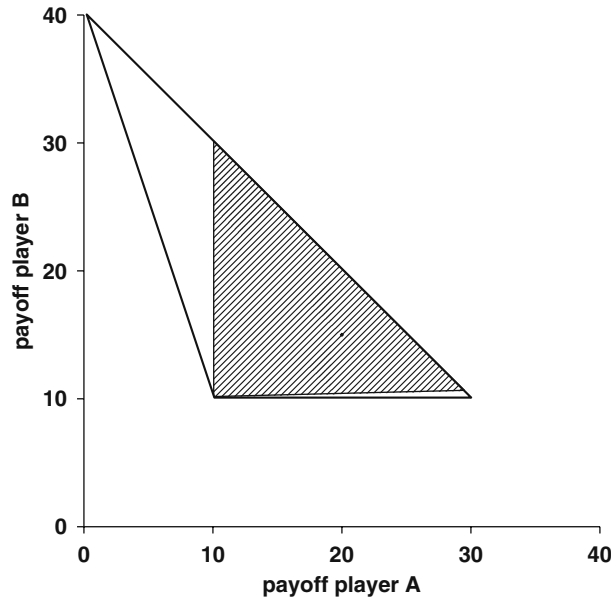
### 2.3. *Evaluating strategies for the investment game*

In this section, we will define the *indefinitely repeated* (iterated) investment game (IIG) and evaluate strategies for this game. The IIG can be defined as follows: The game consists of two players, A and B. After they play the IG once, constituting one period, a new period follows with a probability of  $\delta$  (here,  $\delta = 0.99$ ). Hence, the probability that the game will last for exactly  $t$  periods is  $\delta^{t-1}(1-\delta)$  and the expected number of periods for one game is therefore  $\sum_{t=0}^{\infty} t\delta^{t-1}(1-\delta)$ , which is a geometrical series that converges to  $1/(1-\delta)$ . For  $\delta = 0.99$  the expected number of periods is 100. In every period, both players receive an endowment of 10. Player A decides what integer percentage (0–100%) of the endowment to invest. As before, any investment is multiplied by 3.0 and then sent to player B who decides whether to return any integer percentage of the trebled investment to player A (none of B's own endowment can be passed to A).

Due to the difference in payoff structures (asymmetric versus symmetric) between the IIG and IPD, it can be expected that the evolved strategies will also differ. In particular, the asymmetry of the IIG imposes a unique problem on player A: player A's investment represents a risky outlay because player B could exploit player A by making no return. Player B cannot make up this exploitation in later periods as she cannot herself initiate cooperation again. Instead it has to be player A who reinitiates cooperation by investing. In contrast, in the IPD either player can compensate exploitation by initiating cooperation. This is a crucial difference. Consequently, it can be predicted that the strategies that enable cooperation for the asymmetric IIG will be *less* cooperative than those for the symmetric IPD. A strategy is less cooperative compared to another strategy, when it has fewer conditions under which it cooperates. For instance, Grim is a less cooperative strategy than TFT, because it only cooperates when the opponent cooperated in *all* previous periods, whereas TFT cooperates whenever the opponent cooperated in the *preceding* period.

The game-theoretic equilibrium prediction for the IIG can be drawn from the *folk theorem* for indefinitely repeated games (Fudenberg and Tirole, 1991). The Folk Theorem requires that an outcome should be "individually rational" for each payoff-maximizing player. Any outcome with an average payoff for player A lower than player A's endowment is individually not rational for player A. Likewise, any outcome that gives player B an average payoff below player B's endowment plus the proportion  $1-\delta$  (here 1%) of the trebled investment is individually not rational for player B: Player B on average would profit more by keeping the whole trebled investment in one single period, even if player A thereafter stops any investment, than by keeping less than 1% of the trebled investment in each period of the game (given a constant non-zero investment rate). Figure 3 shows all possible payoff combinations for the game and indicates the payoff combinations that represent equilibria according to the Folk Theorem (the small "gap" at the bottom of the triangle corresponds to the payoffs smaller than 1% of the trebled investment above player B's





*Figure 3.* Payoff region for the indefinitely repeated IG and the prediction of the Folk theorem for indefinitely repeated games. The large triangle indicates all possible payoff combinations for the two players (payoff region). The hatched triangle marks the payoff combinations that are equilibria for the repeated IG with a continuation probability of 0.99, as predicted by the Folk theorem. The diagonal between the two coordinates (30, 10) and (0, 40) represents efficient outcomes that maximize the mutual payoffs.

endowment that player B does not accept). Some equilibria consist of efficient outcomes that maximize the sum of both player's payoffs (i.e. a sum of payoff of 40). For the following, we consider an efficient outcome that leads to a joint payoff of 40 as a "cooperative outcome", whereas any inefficient outcome is considered an "uncooperative outcome". Whether or not a cooperative outcome is obtained depends solely on player A's decision, which has to be an investment of 100%. Unfortunately, as can be seen in Figure 3, the Folk Theorem predicts many equilibria and it is not clear how they are reached. Therefore it does not help us find cooperative strategies for the IIG. Instead, we must turn to consider specific strategies for the repeated game and how they perform. We will represent strategies with finite automata, which can

capture a wide range of strategies, including all the ones discussed above for the IPD.

An automaton can be visualized as a graph of connected states. The automaton starts in the initial state. Given input corresponding to the opponent's last decision, the automaton can move to a different state. For the purpose of parsimony, we assume that each automaton has an "aspiration level" with which the opponents' level of investment or return is sorted into two categories (i.e., player A's move is labeled "Trust" versus "Distrust" for the opponent player B automata, and player B's move is labeled "Reciprocity" versus "Exploitation" for player A automata), producing a dichotomous input. Each state determines an output, which is defined as the player's investment or return rate. Because player B can take player A's first decision (trust or distrust) into account in the first period, the automaton for player B can have two initial states. For the special case in which player A invests nothing, player B can make a decision (e.g., choose a return rate of 40%), but (because any percent of the zero investment is zero) nothing is returned to player A. In this case, player A does not find out anything about player B's decision. Thus following a zero investment, player A's next move cannot depend on player B's decision, and so player A's automata can only move to a single next state.

With this notation, strategies for the IIG can easily be represented. However, the possible number of automata is quite large, even if one restricts the automata considered to a maximum of two states. Therefore we focus our analytical exploration (though not our simulations) on a small but diverse set of just six strategies for player A (Figure 4). The *Never-Invest* strategy never makes an investment, whereas *Always-Invest* always makes an investment of 100%. *Min-Grim* and *Fair-Grim* are Grim strategies that invest fully as long as B's return is above the aspiration levels of 34 and 67% respectively. (The label "Min" refers to the "minimal" return that player B could make—34%—and still provide player A with a barely profitable payoff just above his endowment of 10. The label "Fair" refers to the return rate from player B of

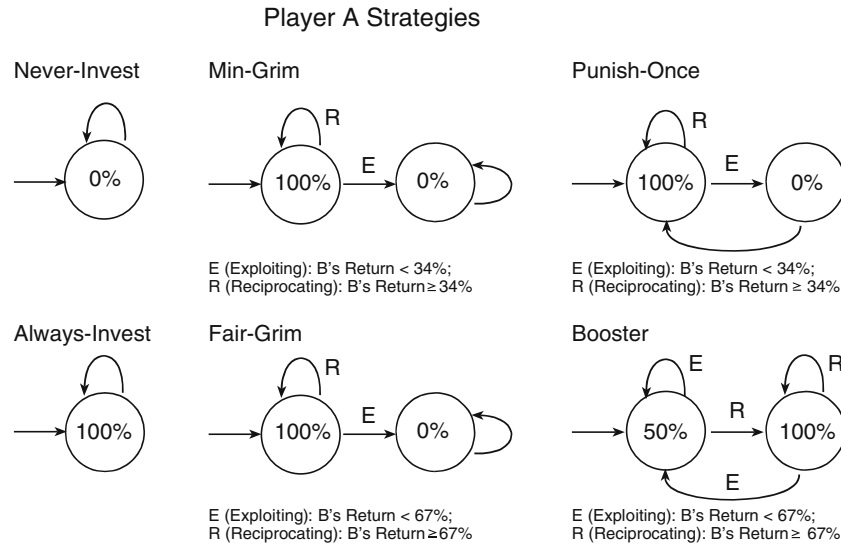
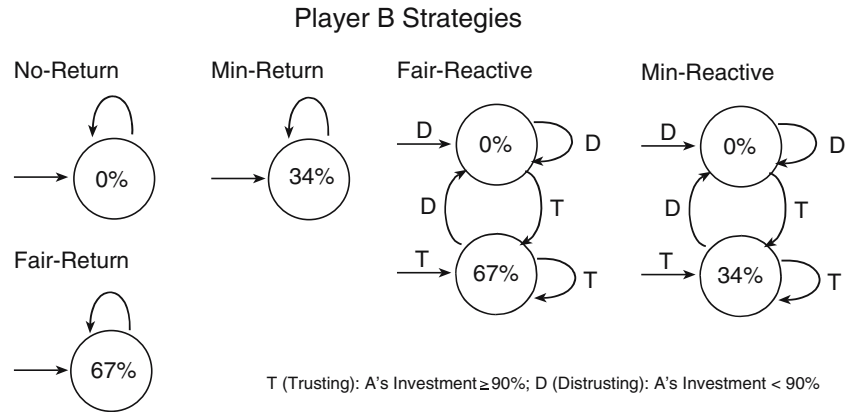


Figure 4. Selected set of strategies for player A. For each automaton states are represented by circles and the initial state is indicated by the left arrow that is not connected to any other state. Each state determines an output, represented by the number inside each circle. Given an input from the opponent, automata may move to other states, indicated by the arrows that start from a state and point to the same or another state; the inputs that cause these particular moves are indicated by the label on each arrow.

67%, which yields an equal “fair” payoff of ~20 for both players if A invests fully.) In contrast to the Grim strategies’ immediate and unforgiving uncooperative reaction to exploitation, the strategies *Punish-Once* and *Booster* are more forgiving. *Punish-Once* repeats an investment of 100% until player B makes a return below the aspiration level of 34%. In this case it makes no investment for one period, thereafter returns to an investment of 100%. *Booster* is an even more tolerant strategy: it starts with an investment of 50%, and then if player B provides a return above or equal to its aspiration level of 67%, it moves to its second state with an investment of 100%. If player B subsequently makes a low return *Booster* moves back to the first state of 50% investment until B gives a high return again.

We restricted player B to a set of five strategies (see Figure 5). Three non-reactive one-state automata for player B



*Figure 5.* Selected set of strategies for player B. For each automaton states are represented by circles and the initial state is indicated by the left arrow that is not connected to any other state. Each state determines an output, represented by the number inside each circle. Given an input from the opponent, automata may move to other states, indicated by the arrows that start from a state and point to the same or another state; the inputs that cause these particular moves are indicated by the label on each arrow.

differ only with respect to their return rates: *No-Return* (0%), *Min-Return* (34%), and *Fair-Return* (67%). The additional two-state automata *Min-Reactive* (returning 34%) and *Fair-Reactive* (67%) are responsive to player A's decision. Both strategies only make their respective returns if player A's investment is at or above their aspiration level, otherwise they return nothing. Their aspiration level was set at 90%, making substantial investments by player A necessary to induce returns by player B.

Although this set of strategies is somewhat arbitrary it includes important characteristics. First, there are basically two crucial investment rates used for the set of strategies: an investment rate of 100% maximizes the possible payoff for player A given player B's behavior, provided B returns at least 34% of the trebled investment. An investment rate of 0% is appropriate if the interaction would lead to a payoff below player A's endowment. Second, there are three crucial return rates for player B: either player B returns nothing to maximize her payoff in a single period, or player B returns 34% to give player A the minimum payoff above player A's

endowment, or player B returns 67%, which leads to equal “putative” fair payoff allocations between both players. To construct the selected set of automata, the crucial investment rates and aspiration levels corresponding to the crucial return rates of player B were selected. In addition, some strategies incorporate a “punishment” mechanism, so that strategies for player A respond to low returns with low investments or, in case of player B, respond to low investments with low returns. In sum, with this small set of strategies it is possible to evaluate strategies’ performance in pairs of strategies A and B combinations, to demonstrate the strengths and weaknesses of particular decision mechanisms. We perform this evaluation in three ways as follows.

### 2.3.1. *Nash equilibrium*

A strategy combination forms a Nash equilibrium if the two strategies are mutual best replies out of the available strategy set.<sup>1</sup> A best reply strategy maximizes the player’s payoff given the opponent’s strategy. Table I shows the expected average payoffs that the strategies reach against each other and indicates the various Nash equilibria. Even with the small set of strategies defined above, many equilibria exist, which lead to a variety of payoff combinations as predicted by the Folk theorem. Player B strategies that consist of only one state can form equilibria that lead to efficient outcomes. In contrast, player A’s one-state automaton that can produce efficient outcomes, Always-Invest, does not form any Nash equilibria. This suggests that, for player A strategies, it may be important to incorporate a “punishment mechanism” that threatens low returns from B. From this it can be predicted that the strategies for player A require a higher degree of complexity to obtain efficient outcomes than do player B strategies, which we will test in Study 1.

### 2.3.2. *Evolutionary stability*

The Nash equilibrium concept for evaluating strategies for the IIG has the problem that the number of strategy combinations

TABLE I  
Payoff matrix for the indefinitely repeated IG for the selected set of strategies

Player A strategies	Player B strategies				
	No-Return	Min-Return	Fair-Return	Min-Reactive	Fair-Reactive
Never-Invest	10.00*, 10.00*	10.00, 10.000*	10.00, 10.00*	10.00, 10.00*	10.00, 10.00*
Fair-Grim	9.90, 10.30	10.01, 10.200	20.10*, 19.90*	10.01, 10.20	20.10*, 19.90*
Min-Grim	9.90, 10.30	10.200*, 29.80*	20.10*, 19.90	10.20*, 29.80*	20.10*, 19.90
Punish-Once	5.00, 25.00	10.200*, 29.80*	20.10*, 19.90	10.20*, 29.80*	20.10*, 19.90
Booster	5.00, 25.00	10.199, 29.70*	20.05, 19.85	5.00, 25.00	5.00, 25.00
Always-Invest	0.00, 40.00*	10.200*, 29.80	20.10*, 19.90	10.20*, 29.80	20.10*, 19.90

*Note:* The cells of the matrix show the expected average payoffs of the strategies for the indefinitely repeated IG, with a continuation probability of 0.99. The first value represents player A's payoff and the second value player B's payoff. The asterisks indicate the best reply strategies for each player. Cells with two asterisks indicate Nash equilibria.

representing equilibria is too large to distinguish the most promising strategies—9 out of 11 strategies in our restricted set appear in Nash equilibria. Thus the concept makes no narrow predictions about which outcomes people are likely to reach and which strategies they may employ in the IIG. One way to address this problem is to explore other plausible and more restrictive approaches for evaluating strategies such as the concept of evolutionary stability (see Bendor and Swistak, 1998; Maynard Smith, 1984; Maynard Smith and Price, 1973; Samuelson, 1997; Weibull, 1995). Because this concept is typically only defined for symmetric games, we transform the asymmetric IIG into a symmetric game (see Samuelson, 1997), by assuming that a strategy for the symmetric game is represented as a combination of a strategy for player A and a strategy for player B. The payoff for a strategy in the symmetric game is defined as the sum of the expected average payoffs of the two strategies of which it consists. An *evolutionarily stable strategy (ESS)* is a best reply against itself, thus representing a Nash equilibrium. Furthermore, the crucial stability criterion is that if an alternative strategy ( $\sigma$ ) leads to the same payoff against the ESS as the ESS reaches against itself, then it is required that the ESS must lead to a greater payoff against the alternative strategy than the alternative strategy reaches against itself:

$$\begin{aligned} \text{payoff}(ESS, ESS) = \text{payoff}(\sigma, ESS) &\rightarrow \text{payoff}(ESS, \sigma) \\ &> \text{payoff}(\sigma, \sigma). \end{aligned}$$

In contrast to the Nash equilibrium concept, which is weak for indefinitely repeated games (i.e., too many strategy combinations fulfill its requirements), the evolutionary stability concept is strong (i.e., its requirements are difficult to fulfill). From our set of six player A and five player B strategies presented above, 30 combined strategies for the symmetric game can be composed. None of these combined strategies are evolutionarily stable, because none fulfill the stability criterion. This result is not surprising, as it has been shown for the IPD that no ESS exists given a substantial continuation probability

(Boyd and Lorberbaum, 1987; Lorberbaum, 1994). Thus the ESS concept does not help us in identifying a smaller set of reasonable strategies for the IIG either.

An alternative concept of *limit evolutionarily stable strategies* (Limit ESS) appears to be more appropriate for evaluating repeated game strategies (Leimar, 1997; Samuelson, 1991; Selten, 1983, 1988). For Nash equilibrium strategies, multiple best reply strategies often exist, disqualifying the equilibrium strategy as an ESS. These ties in payoffs can often be broken if players in the game make errors with a small probability, which can be called a “perturbed game” (Selten, 1983). A strategy is called a Limit ESS for a game (in which no errors occur) if it is an ESS for the perturbed game in which errors do occur (Selten, 1983, p. 304). Thus, to identify the Limit ESSs, one simply has to incorporate small errors in a game and identify the ESSs for this game with errors.

When playing an indefinitely repeated game “strategy selection” errors could occur, which implies that sometimes at the beginning of a game (i.e. in the first period of the repeated game) a player selects an unintended strategy. For exploring whether the above set of strategies contains Limit ESSs when selection errors occur, we assume that a player will accidentally select an unintended strategy (at random) in 1% of the games. Table II shows the corresponding payoffs for each strategy if both players select their strategies with such small errors (the strategy’s payoff can be directly calculated as the payoff it obtains against the strategy with which it is paired multiplied by 0.99, plus each payoff it obtains against all alternative strategies it could erroneously be paired with multiplied by  $0.1/n$ , where  $n$  is the number of alternative strategies). Thus in contrast to the ESS concept, which explores the stability of a strategy combination compared to each alternative strategy individually, the Limit ESS concept explores the stability of a strategy combination compared to the whole set of alternative strategies simultaneously. When selection errors occur, the Min-Grim and Min-Return strategy combination represents an ESS for the symmetric version of the IIG, and thus a Limit ESS for the IIG without selection errors.



TABLE II  
Payoff matrix for the perturbed indefinitely repeated IG with selection errors

Player A strategies	Player B strategies				
	No-Return	Min-Return	Fair-Return	Min-Reactive	Fair-Reactive
Never-Invest	9.96*, 10.12	10.00, 10.10	10.10, 10.10	9.99, 10.15*	10.07, 10.11
Fair-Grim	9.91, 10.46	10.06, 10.40	20.00*, 19.81	10.05, 10.39	19.97, 19.82*
Min-Grim	9.91, 10.56	10.25*, 29.62*	20*, 19.9	10.24*, 29.61	19.98*, 19.91
Punish-Once	5.12, 24.94	10.24, 29.66*	19.99, 19.94	10.23, 29.65	19.96, 19.95
Booster	5.07, 24.94	10.18, 29.56*	19.89, 19.89	5.10, 24.97	5.18, 24.93
Always-Invest	0.23, 39.61*	10.22, 29.69	19.98, 19.98	10.21, 29.69	19.95, 19.99

*Note:* The cells of the matrix show the average payoff the strategies obtained against each other, if players make selections errors with probability 0.01.

Therefore the Limit ESS concept, contrary to the Nash equilibrium and ESS concepts, indicates that one strategy combination from the selected set outperforms the others in terms of their stability under the occurrence of selection errors. These low-probability errors have the effect of breaking ties that were previously present between strategies' payoffs. As can be seen in Table 1, when no errors occur, the strategies Min-Grim, Punish-Once, and Always-Invest for player A perform equally against four of player B's strategies. These ties can be broken in favor of Min-Grim and Min-Return, when small selection errors occur (see Table II).

The assumption that people both commit errors with low probability and also take into account the possibility that others will commit errors when making strategic decisions is psychologically very reasonable. Therefore it is plausible to use the Limit ESS concept to evaluate strategies: Strategies have to be best replies and have to be stable under the presence of errors such that other strategies cannot replace them. Our theoretical analysis demonstrated that the many strategies picked out by the Nash equilibrium concept could be reduced to one single strategy combination when the additional stability criterion is imposed. However, we hesitate to draw general conclusions about the strategies' characteristics, because we started with a small selected set of strategies initially. Instead, we next turn to using evolutionary simulations to greatly extend the set of strategies that can be considered in our search for those that may underlie cooperative behavior in the IIG.

### 3. STUDY 1: COOPERATIVE STRATEGIES IN THE IIG

To find out what kinds of strategies can produce cooperation, that is, trust and reciprocity, in an asymmetric social situation, we simulated in Study 1 an evolutionary process for the IIG. In this simulation, a population of agents, equipped with different strategies, played the IIG against each other. In addition, for comparison, we ran a similar simulation for the IPD (with the payoff matrix defined in Figure 2). These simulations follow

the Limit ESS concept as they incorporate the idea of selection errors discussed above: during the evolutionary process, agents occasionally enter the population using new or altered strategies that differ from those already in common use. Due to the difference in payoff structures (asymmetric versus symmetric) between the IIG and IPD, it can be predicted that the strategies that enable cooperation for the asymmetric IIG will be *less* cooperative than those for the symmetric IPD. Specifically, we can ask whether the less-cooperative Min-Grim and Min-Return strategy combination identified above will also frequently appear in the evolutionary simulations.

### 3.1. *Method*

#### 3.1.1. *Representing strategies*

The finite automata for strategies were represented as vectors. Each automaton's state requires three elements of the vector, one for the output and two to identify the next states that can be moved to depending on the dichotomous input. For the IIG, the output was restricted to multiples of 10 ranging from 0 to 100 for simplification. Automata for the IPD have two possible outputs, "Cooperate" or "Defect". In addition, automata for the IIG have an aspiration level ranging from 0 to 100 for categorizing the opponent's decision as "Trust" or "Distrust" for player A's investments or as "Reciprocity" or "Exploitation" for player B's returns. Each automaton in the IIG or the IPD also specifies an initial state, and all automata were restricted to a maximum of two states; thus two-state IIG automata were represented by an eight-element vector and those for the IPD used seven elements.

#### 3.1.2. *Simulating an evolutionary process*

To simulate the evolutionary process, we used a genetic algorithm (Goldberg, 1989; Michalewicz, 1996; Mitchell, 1996) as is becoming common in evolutionary game theory (e.g. Axelrod, 1987; Hoffmann, 1999, 2001). Agents compete against each other in pairs. In the first generation, for each of the 100 agents

in the IIG two automata were randomly generated, one for the player A role and one for the player B role, whereas in the IPD only one automaton was needed for each agent. The automata outputs, states, and aspiration levels were drawn with equal probability from the set of possible values. At every generation, each agent in the IIG played 50 games in each role (100 games total), and each agent in the IPD played 50 games, with a continuation probability  $p = 0.99$  at each period. For each game, an opponent agent was drawn randomly with the constraint that an agent never played against itself or twice against another agent. Each agent's fitness was defined as the average per-period payoff (averaged across the different lengths of the games) that the agent obtained for all games it played.

Based on their fitness, agents were selected for the next generation via a tournament selection procedure (see Goldberg and Deb, 1991; Michalewicy, 1996). Each agent took part in six different "tournaments" (pooled comparisons) of six randomly-chosen agents. From each of the 100 tournaments, the agent with the highest fitness (averaged over the two roles in the IIG) was selected for the next generation. If multiple agents tied for the highest fitness, one was randomly selected. To create population variation after the selection procedure, all agent automata were randomly put together in pairs, and then with a probability of  $p = 0.80$  each pair had some of their vector elements swapped via a two-point crossover procedure (separately for player A and player B in case of the IIG) as follows: if a crossover occurred, two positions of the vectors representing the automata were determined randomly with equal probability, and all vector elements between these two positions were interchanged between the two "parent" strategies. After the crossover procedure, automata were further modified through a mutation procedure that could alter each vector element with a certain probability. This mutation probability depended on the automata's number of states such that there was an overall probability of  $p = 0.33$  that at least one element of the automaton was mutated (e.g., for an automaton with two states for player A in the IIG this implied a per-element mutation probability of  $p = 0.049$ ). If

a mutation occurred, a new output value or state was drawn with equal probability from the set of possible outputs and states. For the aspiration level of the automata for the IIG, new values were drawn from a normal distribution with the mean equal to the old aspiration level<sup>2</sup> and a standard deviation of 5% points.

This genetic algorithm was run 100 times, each for 1000 generations. Since automata with different numbers of states can still be equivalent in terms of their behavioral output, after every 10 generations all evolved automata (each having a maximum of two states) were transformed to their corresponding minimal automaton, that is, the automaton out of the set of equivalent automata with the minimum number of states (Hopcroft and Ullman, 1979). The last (1000th) generation was analyzed as a snapshot of where the ongoing evolutionary process was headed. Because of mutation and crossover, the last generation always contains at least a few low-performing strategies in addition to the high-performing ones of interest. These ineffective strategies were screened out by continuing the evolutionary selection process alone for 200 further generations (without crossover or mutation), yielding the final population of strategies that we evaluated.

## 3.2. *Results*

### 3.2.1. *Characteristics of the evolutionary process*

The evolutionary dynamics varied substantially across the different runs for the IIG, but some general patterns can be identified. In the beginning of the evolutionary process, the populations quickly evolved to one-state automata for the IIG: Never-Invest for player A and No-Return for player B, which together lead to a payoff of around 10 for each player. This configuration is often kept for long periods (hundreds of generations) until the population suddenly changes, frequently switching to a Min-Grim and Min-Return strategy combination. This new pattern, providing efficient payoffs of around 12 and 28 for A and B, respectively, can also remain stable for a

long period. Following this, another population transition may lead back to the Never-Invest and No-Return strategy combination, and so on. Figure 6 shows the average payoffs obtained in one such run of the evolutionary process. During the first 450 generations, the payoff for each player is around 10, after which payoffs increase to 12 and 28 when the population evolves to the Min-Grim and Min-Return strategies.

For the IPD, the picture is different. At the beginning of the evolutionary process Always-Defect takes over the whole population, leading to payoffs of 10. In all runs this configuration does not change substantially; in other words, in none of the runs did the evolutionary process lead to the development of cooperative strategies producing efficient outcomes.<sup>3</sup>

### 3.2.2. *Strategies evolved over the course of evolution*

Because of the variation across runs for both the IPD and IIG, we now turn to an analysis of the evolved automata of the last (1000th) generation. Since the strategy population in the 1000th generation is the result of an evolutionary process of one single run, each population is independent of all other populations, which is beneficial for statistical comparison. In 81 out of the 100 runs for the IIG, the *predominant strategies*, defined as those strategies that were applied by more than 50% of the agents in a population, in the last generation led to inefficient outcomes (payoffs of 10 for both players). In the other 19 runs, the predominant strategies led to efficient payoffs of around 12 for player A and 28 for player B. Due to the restriction of investment and return rates to multiples of 10%, the minimum amount above the original endowment that player A could obtain was 12.

There was clear convergence on the strategies used in the efficient and inefficient IIG outcome cases, respectively. In all 81 runs with inefficient outcomes, the agents always used Never-Invest and No-Return. For each of the 19 runs with efficient outcomes, the predominant strategy for player A was Min-Grim with an average aspiration level of 33% (SD = 4.4%). In 14 out of the 19 runs, the Min-Grim strategies were

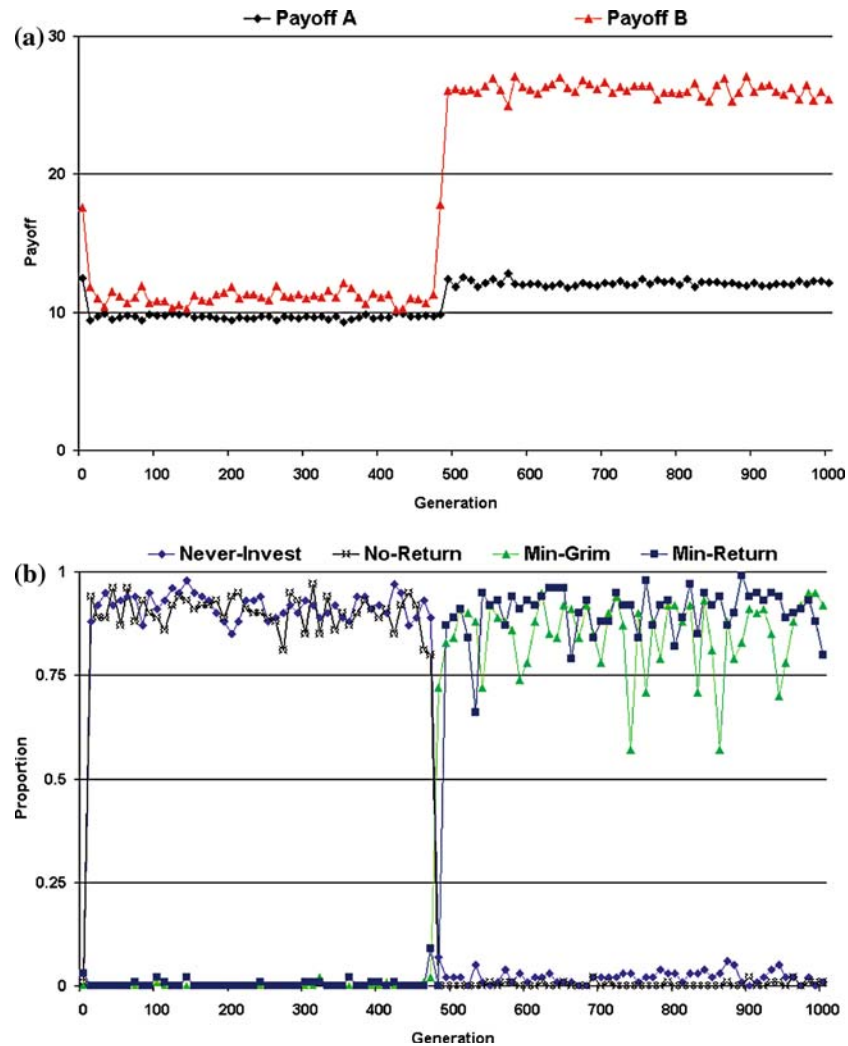


Figure 6. An example run of the simulated evolutionary process for the repeated IG in study 1. Figure 6A shows the strategies' average payoffs for player A and player B. Figure 6B shows the proportion of the Never-Invest and Min-Grim strategies for player A and the No-Return and Min-Return strategies for player B in the population.

paired with the Min-Return strategy for player B, while the other five runs had different two-state automata for player B. The behavior of these alternative strategies is quite similar to Min-Return as they always return 40% of the trebled

investment when a substantial investment is made by player A, and only in the case of very low investments do these strategies move to a second state with lower returns. These results mostly match the strategies found with the Limit ESS concept earlier.

For the IPD, the evolutionary process converged strongly. In all 100 out of the 100 runs, the population of agents applied the Always-Defect strategy, leading to an inefficient outcome with payoffs of 10. Thus, the proportion of efficient outcomes of 19% for the IIG is significantly larger than the proportion of 0% efficient outcomes in the IPD [ $\chi^2(1, N = 200) = 20.99, p = .001$ ; corresponding to a large effect size of  $h = 0.90$  according to Cohen, 1988].

### 3.3. *Summary of Study 1*

Study 1 demonstrates that the strategies that most frequently evolve for the IIG and the IPD lead to inefficient outcomes. The strategies that produced the inefficient outcomes for both games were very similar as they consist of only one “uncooperative” state: the Never-Invest and No-Return strategies for the IIG, and the Always-Defect strategy for the IPD.

The important difference between the games is that for the IIG a substantial proportion of runs were observed where the evolutionary process led to a population of cooperative strategies in the last generation, whereas for the IPD the evolutionary process never led to cooperative strategies. This is surprising as one would expect cooperative strategies like TFT to evolve, considering the results of previous studies (e.g. Axelrod, 1984).

Although for the IIG the proportion of the evolutionary processes that led to efficient outcomes was much larger than for the IPD, the strategies that enabled the efficient outcomes for the IIG were not very cooperative: the Min-Grim strategy for player A starts with a cooperative decision, but the strategy responds to a single low return from player B by switching to no investment for all following periods, eliminating the possibility for any further cooperation. Thus Min-Grim



is an unforgiving strategy and therefore it is less cooperative than strategies such as TFT, which always allows a return to a cooperative state at the opponent's initiative. Furthermore, even when the evolutionary process led to an efficient outcome in the IIG, the payoff allocation between players A and B was extremely skewed in B's favor. Equal "putative" fair payoff allocations were not observed.

#### 4. STUDY 2: INCREASING STRATEGY COMPLEXITY

In Study 2 we explored the impact of allowing greater potential strategy complexity in the two games. In Study 1, the automata representing the strategies were restricted to a maximum number of two states, yielding strategies that are rather simple.<sup>4</sup> For the IPD it has been shown that increased potential complexity of strategies could alter what evolves, such as strategies that take mistakes of the opponent into account (Lindgren, 1991) or are otherwise able to exploit opponent strategies (Axelrod, 1987). Therefore it is reasonable to conclude that the restriction of strategy complexity to two states in Study 1 could have promoted the emergence of simple strategies like Min-Grim. These strategies might be outperformed by more sophisticated and "complex" strategies when the restrictions on the number of states is relaxed. In addition, the restriction to strategies with a maximum of two states might also have prevented the emergence of cooperative strategies for the IPD that require more than two states (unlike TFT, which requires only two states). Thus, in Study 2 we aimed to find out whether more memory (and hence more information) along with more computation could allow the more frequent evolution of cooperative strategies.

##### 4.1. *Method*

We made two changes to the simulations used in Study 1. First, the maximum possible number of states for a strategy was doubled from two to four. Second, to have the same

overall probability of a mutation occurring in an automaton, the per-element mutation rate had to be reduced (i.e. for a four-state automaton the mutation probability had to be  $p = 0.028$  for each parameter in the strategy).

## 4.2. Results

As in Study 1, the evolutionary dynamics differed substantially across runs. Here we will report only the predominant strategies in the last generation of the evolutionary process with 1000 generations.

### 4.2.1. Strategies evolved for the IIG

For the IIG, the increase in the number of allowed states did not change the outcome of the evolutionary process substantially. In 71 out of 100 runs (compared to 81 in Study 1), the evolutionary process led to an inefficient result with a payoff of around 10 for both players in the last generation. For all of these inefficient outcomes, player A used Never-Invest. For player B, the strategy No-Return was found in 55 of these 71 runs, with variants of No-Return in the remaining 16. These variants, despite consisting of more than one state, did not differ substantially from the No-Return strategy: all of them did not make a return if a substantial investment was made, but increased their return if they received a low investment.

Efficient outcomes (with payoffs 12 and 28 for A and B) arose in 23 out of the 100 runs. Min-Grim was used by player A in all these 23 runs, with a mean aspiration level of 34% ( $SD = 4\%$ ). For player B, Min-Return was used in 11 out of 23 runs; in the other 12 runs, strategies with a higher number of states were applied. However, similar to Min-Return, these strategies always make the “minimal” return of 40% when player A makes a substantial investment (giving player A the minimum payoff above his endowment, i.e. 12), but make no or very small returns if a very low investment was made. In sum, when the evolutionary process led to an efficient outcome, the evolved strategies were very similar to the

Min-Grim and Min-Return strategy combination found in Study 1. In the remaining six of the 100 runs almost efficient outcomes were produced with an average payoff of 33 for both players (12 for player A and 21 for player B). However, the populations consisted of rather mixed strategies.

In contrast to the IIG, the increase of potential strategy complexity had a strong impact on the evolutionary outcomes for the IPD. The number of inefficient outcomes in the last generation arising from predominant use of Always-Defect decreased to 74 out of 100 runs (compared to 100 inefficient runs in Study 1). In the remaining 26 runs, the strategies almost produced efficient outcomes with mean payoffs of around 19 ( $SD = 1.6$ ). The strategy TFT was not observed in any of these near-efficient runs. Instead in 22 runs the strategies consist of four states and in four runs the strategies consist of three states. Altogether 18 distinct strategies were observed, all sharing some similarities: most of them (16) always defect in the first period. Seven of them stay in the first state and continue defecting as long as the other strategy cooperates, thereby exploiting unconditional cooperative strategies. Fourteen defect in the second period when the other strategy also defected in the first period. However, if the opponent strategy defects in the first and the second period 7 of the strategies cooperate in the third period. Moreover, 15 of the 18 strategies entail a cooperative state like TFT, so that the strategy cooperates as long as the opponent cooperated in the last period. Only two of the 18 strategy entail a terminal defecting state like all Grim strategies, in which the strategy continues to defect until the end of the game regardless of the opponents' decision. To illustrate one of the strategies, Figure 7 shows the strategy that was observed most often in the last generation of the evolutionary processes (in 4 out of 22 efficient runs), labeled *Cautious TFT*. The strategy behaves very similarly to the standard TFT strategy, with the difference that it begins by defecting in the first period. Then it always cooperates if the opponent cooperated in the previous period. After a maximum of four defecting decisions of the opponent the strategy will cooperate for at least one round.

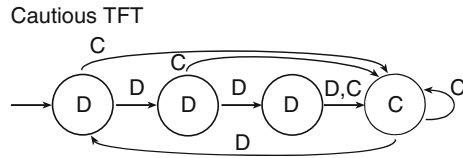


Figure 7. A strategy, named cautious TFT, which was predominant in the efficient runs for the indefinitely repeated PDG in Study 2.

Overall, the 29% proportion of runs with efficient outcomes for the IIG did not differ significantly from the 26% proportion of efficient outcomes for the IPG [ $\chi^2(1, N = 200) = 0.23, p = .64; h = 0.07$ , which is smaller than what is commonly considered as a small effect size according to Cohen, 1988].

#### 4.3. Summary of Study 2

We found that increasing the potential complexity of strategies increased the frequency of efficient outcomes for the IPD, but had little effect for the asymmetric IIG. On the one hand the proportion of IIG runs where the strategies in the 1000th generation produced efficient outcomes did not differ from the proportion observed in Study 1. On the other hand the IIG strategies that evolved did not differ substantially from the strategies observed in Study 1: basically the same strategies produced the efficient and the inefficient outcomes.

The larger potential complexity of the strategies had a strong effect on the evolutionary process for the IPD, as a substantial proportion of efficient outcomes were observed in the last generation. The strategies that produced the cooperative outcomes differ from TFT, consistent with previous results (Axelrod, 1987; Lindgren, 1991), mostly by starting with an uncooperative decision in the first period. Some of them can thereby exploit unconditionally cooperative strategies that cooperate regardless of whether the other strategy cooperates or defects. However, most strategies also include a mechanism to establish cooperation, so that after some periods the strategies cooperate when the opponent strategy also cooperates. Thus, the increased potential complexity of the strategies allowed them to incorporate mechanisms both

to exploit others and to establish mutual cooperation. These results suggest that in a symmetric bargaining situation where exploitation is possible, increased strategy complexity may be an advantage because it enables more devious exploits. On the other hand, in an asymmetric bargaining situation like the IIG, where player A has no possibility to exploit the opponent, greater strategy complexity may be of no use.

## 5. STUDY 3: DIFFERENCES BETWEEN THE IIG AND THE IPD

Studies 1 and 2 show that there are substantial differences between the kinds of evolved strategies that produce efficient outcomes for the IIG and the IPD. What causes these differences? One distinction between the two games is that the two players make decisions sequentially in the IIG, whereas they make decisions simultaneously in the IPD. A second distinction is that the players in the IIG have a continuum of possible decisions, whereas they make dichotomous decisions in the IPD. A third distinction between the IIG and the IPD is of course that the payoff matrix is asymmetric for the former and symmetric for the latter. To test whether these distinctions cause differences in the evolved cooperative strategies, we modified the IIG by removing the differences between the games in two steps. In the first condition the IIG was modified such that the players had to choose between two dichotomous decisions simultaneously. In the second condition the IIG was further modified to remove the asymmetry of the payoff matrix. The interesting question is what distinction causes a different evolutionary process for the IIG compared with the IPD.

### 5.1. *Method*

In Study 3 we conducted simulations in two distinct conditions. In the first *simultaneous and dichotomous IIG* condition, we ran an evolutionary process for the IIG as done in Study 1, but with two important changes: both players had to make

their decisions simultaneously. This has a major impact on player B's decision process. Whereas in the sequential game, player B could take player A's decision (investment) in the present period into account when making a return decision, in the simultaneous game player B (like player A) can only consider decisions of previous periods. Therefore, in these simulations the automata for player B could only change their states (if at all) in response to player A's investments in the previous period, without knowing player A's decision in the present period. In addition we changed the IIG so that both players only have two decision alternatives. Player A can either invest nothing or his entire endowment (i.e. a 100% investment) and Player B can either return nothing or 34% of the investment. If player A invests nothing both players receive a payoff of 10. In contrast, if player A invests all, the two players' payoffs depend on player B's decision: If B returns nothing A earns nothing and B receives a payoff of 40. If B returns 34%, A receives 12 and B receives 28.

In the second *symmetric IIG* condition we ran an evolutionary process for the IIG as in the simultaneous and dichotomous condition, but with one additional change to the modified IIG: if player A decides to make no investment and player B decides to make a return, then A receives a payoff of 40 and B receives 0. This makes the game almost identical to the IPD represented in Figure 2, with the remaining difference that if both players "cooperate", that is when player A invests and player B makes a return, player A receives 12 and player B receives a payoff of 28 instead of an equal payoff of 20 like in the IPD. However, if the payoffs of both games were transformed into ranks according to the player's payoff, this modified IIG becomes identical with the IPD.

## 5.2. Results

In the simultaneous and dichotomous IIG condition the strategies of the population in the last generation produced efficient outcomes in 14 of 100 runs. This proportion does not differ significantly from the 19% proportion of efficient

outcomes observed for the IIG in study 1 [ $\chi^2(1, N = 200) = 0.91, p = .45; h = 0.14$ , which is lower than a small effect size according to Cohen, 1988]. Strategies similar to those in Study 1 were observed: of the 86 inefficient runs, for player A the Never-Invest strategy was observed in 85 runs and for player B the No-Return strategy was observed in 68 runs. Likewise for the 14 efficient runs, in 13 runs the Min-Grim strategy was observed for player A and in six runs the Min-Return strategy was observed for player B. In the remaining runs variants of the Min-Return strategy were observed for player B, which make a return in the first period and continue doing so as long as the opponent makes an investment.

When we also made the IIG symmetric, along with dichotomous and simultaneous decisions, the proportion of efficient outcomes in the last generation of the evolutionary process decreased substantially, so that in none of the populations in the last generation were efficient outcomes observed. This proportion differs substantially from the proportion of 19% efficient outcomes observed for the IIG in Study 1 [ $\chi^2(1, N = 200) = 20.99, p = .001; h = 0.90$ , corresponding to a large effect size according to Cohen, 1988]. In all 100 runs the strategies Never-Invest and No-Return produced the inefficient outcomes.

### 5.3. *Summary of Study 3*

In the simultaneous and dichotomous condition of Study 3 the proportion of efficient outcomes was similar to the proportion of efficient outcomes in Study 1. Moreover the strategies that evolved for the dichotomous and simultaneous IIG do not differ from the observed strategies in the sequential game with continuous decision alternatives. In contrast, when the asymmetry is also removed from the game by making the ranks of the player's payoffs for the IIG identical with the ranks of players' payoffs in the IPD, the evolutionary process of the modified IIG becomes non-distinguishable from the evolutionary process of the IPD. Therefore we can conclude that the different cooperative strategies observed for the IIG compared to the IPD are mainly due to the difference

in the payoff structure of the two games rather than to the sequential decision process or the continuous decision alternatives of the IIG.

In the first condition with the modified simultaneous and dichotomous form of the IIG, the risk for player A of trusting player B is similar to the risk of cooperating in the IPD, namely, exploitation by the other player. However, the opportunities for overcoming this risk differ in the two games because of the difference in symmetric versus asymmetric payoffs. Whereas either player who exploits the other in the IPD can compensate the exploitation through later cooperation, this compensation is not possible in the IIG: player B, who can exploit, cannot “make peace” unless player A tries to do so first through another exploitable investment. When we allow player B to make such a conciliatory gesture in the “symmetric IIG” condition, “returning” 40 to A when A has invested nothing, this allows a cooperative situation to reemerge, showing the benefit of symmetric payoffs for reconciliation.

## 6. DISCUSSION

In this paper, we have asked whether individuals’ cooperative decision strategies in an asymmetric social relationships represented by the IIG differ from those used in symmetric relationships represented by the IPD. This question is particularly important because of the prevalence of asymmetric social relationships: individuals’ positions in relationships are often hierarchical and not interchangeable (remember the employer and employee in the introduction). If asymmetry has an impact on decision strategies, then the predominant use of the symmetric IPD as a general model for social interactions is not justified, and an extension to models using asymmetric games like the IIG appears necessary. The results of the studies reported here support just such a conclusion.



### 6.1. *Game-theoretic concepts*

We have illustrated how three different game-theoretic concepts can be used to evaluate repeated game strategies. Even with a small selected set of strategies we could show that the Nash equilibrium concept is too weak because too many strategies represent equilibria. One way to tackle the resulting equilibrium selection problem is to require additional abilities of the strategies. An important advantage for any strategy is persistence in the long run, evolutionarily speaking. This idea, captured in the ESS concept, forms a persistent population that cannot be “invaded” by any alternative strategy. However, the ESS concept is too strong: among our selected set of strategies for the IIG no uninvadable ESS could be found. The third concept of Limit ESS (Selten, 1983, 1988) is based on the reasonable assumption that people make small errors in their decision processes. Using this concept, we were able to distinguish a strategy combination that is an evolutionary stable strategy when errors occur, namely Min-Grim with Min-Return, which together allow cooperation.

### 6.2. *Evolutionary simulations*

Only a limited number of strategies can readily be assessed in terms of game-theoretic equilibrium concepts. With evolutionary simulations, however, a large set of strategies can be evaluated. In Study 1, where automata were restricted to a maximum of two states, the evolutionary process typically led to two outcomes for the IIG. First and most frequently, inefficient outcomes were obtained by the strategies Never-Invest and No-Return. Second, the less common efficient outcomes were usually associated with Min-Grim and Min-Return. The IPD differed from the IIG as only inefficient outcomes were obtained in the last generation, using the always-defect strategy.

What is the reason for the frequent emergence of the Min-Grim strategy in the IIG? One strength of Min-Grim is that it can only be exploited once, after which it puts itself in

an unexploitable position by no longer investing. However, this advantage is also a disadvantage, because it disables Min-Grim from ever obtaining efficient outcomes again once it has moved to its inefficient non-investing decision (i.e. its second state). Therefore, the Min-Grim strategy, while being as simple as TFT, is less cooperative. The crucial aspect of the IIG seems to be that when player B defects and exploits A, it is only possible to reestablish cooperative outcomes again if player A—the exploited, not the exploiter—makes a risky outlay. In this respect, the asymmetric IIG differs from the IPD. In the IPD, a player that defects and causes an uncooperative situation can reinitiate cooperation through a cooperative decision in a subsequent period. This cooperative gesture will give the player a lower payoff than the other player if the other defects. In this way, a gain by an exploitative decision can be compensated by creating the possibility for a balancing loss (and relative gain for the other player) in another period. In contrast, in the IIG the gain from an exploitative decision by player B is not compensated by a loss for player B when returning to a cooperative interaction. Similarly, the loss for player A resulting from player B's exploitation will not be compensated when the players again cooperate. Furthermore, in order to return to the cooperative interaction, player A has to again make a high investment and bear the risk of a repeated exploitation and hence additional loss. In sum, this burden of trust being placed entirely on one player (A) as a consequence of the asymmetric payoff structure seems to be the crucial difference between the IIG and the IPD.

Player A in the IIG thus faces a dilemma: He has to avoid being exploited, yet he can only increase his payoff through substantial investment, which would again make him vulnerable to exploitation. In contrast, the strategic situation for player B appears to be quite simple: She merely has to make a decision about the rate of returns. Promoting high investment does not seem to be particularly important, which explains the low variance of the strategies obtained for player B across all studies.

Study 2 showed that allowing greater strategy complexity had no substantial effect on the strategies that evolved for the IIG, though it did enable the evolution of more complex strategies for the IPD, which led to efficient outcomes. When the complexity of the strategies was increased, the proportion of evolutionary processes that led to efficient outcomes was the same for the IIG and the IPD. However, the TFT strategy was not observed in any evolutionary process of the IPD. This might be because the simple TFT strategy for the IPD is not only vulnerable to mistakes (Lindgren, 1991), but also unable to exploit naïve, cooperative strategies. The more complex IPD strategies that evolved in Study 2 are capable of both cooperating with other strategies and exploiting them when possible. In contrast, for the asymmetric IIG, extra complexity does not help player A to cooperate without being exploited, nor does it help B to figure out when she is playing with an overly-cooperative A who can be exploited without fear of perpetual defection.

Study 3 demonstrated that the sequential decisions and the continuous decision alternatives of the IIG do not affect the kind of strategies that evolve for the game. This could be explained by the fact that the order in which the decisions are made should only be relevant for player B (because B may or may not take A's decision into account). But because the strategies that typically evolve for player B are very simple (No-Return or Min-Return, with only one state) and do not take into account *any* move on A's part, it actually makes no difference to them whether decisions are made sequentially or simultaneously. However, when the asymmetric payoff structure is removed from the IIG, the evolutionary process changes dramatically and becomes indistinguishable from the evolutionary process of the IPD. Thus, it is the asymmetry of the decision situation, rather than its sequentiality or continuous decision alternatives, that makes the IIG a unique challenge, distinct from the IPD.

### 6.3. *Fair allocations*

In almost all simulations for the IIG in which an efficient outcome evolved, the players obtained unequal payoffs. Player A only obtained a little more (12) than his endowment whereas player B got almost all (28) of the produced surplus. This result was expected from our theoretical analysis because the Fair-Grim and Fair-Return strategy combination (with equal payoffs) is weakly dominated by the Min-Grim and Min-Return strategy combination (with unequal payoffs), receiving a lower or equal total payoff. The reason for this is that Grim strategies for player A that apply a low aspiration level for classifying player B's return as "reciprocating" earn a higher payoff than other Grim strategies with a higher aspiration level, because the former cooperate with a wider range of player B strategies. In other words, Grim strategies with a lower aspiration level than Fair-Grim "underbid" that strategy and thereby undermine the "norm" of high returns that Fair-Grim promotes.

This process resembles a common good problem (see Ledyard, 1995) in which the common good is represented by high returns of player B. If players A restrict themselves to taking only high returns (like Fair-Grim does), the common good of high returns for player A will be maintained in the long run; therefore, this self-restriction is socially preferable. However, for each player A, it is individually rational also to accept low returns in every period, rather than ending the interaction with the current B partner and getting nothing. If a large number of A players make this individually rational decision, it also becomes individually rational for each player B to lower her return rate, and the original equitable division unravels to unequal payoffs.

### 6.4. *Experimental evidence*

How do the results of our evolutionary simulations relate to experimental evidence? There is a growing body of experimental findings on behavior in asymmetric games like the

*ultimatum game* (Güth, et al., 1982; Ochs and Roth, 1989), or the *trust game* (Güth et al., 1997). In the sequential “one-shot” trust game each player has two choice alternatives. The first player can distrust the second player, which leads to a medium payoff for both players. Or the first player can trust the second, which produces a larger payoff for both players if the second player reciprocates the trusting decision. However, if the second player exploits the first’s trust, the first player earns nothing and the second player receives the highest payoff in the game. The majority of participants (82%) in the role of the first player chose not to trust their opponent, and the majority of participants (71%) in the role of the second player chose to exploit their opponent, which is consistent with the game-theoretical prediction. The relatively low magnitude of cooperation might be explained by the fact that the game was not repeated.

Rieskamp and Gigerenzer (2005) studied the IIG experimentally and observed a large magnitude of cooperation: the average mutual payoff received was 34 compared to the possible maximum of 40. However, this cooperation did not lead to equal payoffs for both players. Consistent with our evolutionary simulations, participants in the role of player B reached a higher average payoff compared to player A participants (19 versus 15). Thus some of the participants realized the power of the player B role and used it to obtain a higher payoff. However, participants often also made decisions that led to equal payoffs for both participants, thereby following fairness principles (Deutsch, 1975). Additionally Rieskamp and Gigerenzer (2005) showed that a few surprisingly simple heuristics could account for individuals’ decisions in the IIG. The strategy that best modeled participants’ decisions in the role of player A resembled the Grim strategy found in our evolutionary simulations. Like Grim, this observed strategy starts with an investment of 100% and maintains this level as long as a return above 34% is made. In contrast to Grim, this strategy stays at 100% investment if player B occasionally makes a low return; but if player B repeatedly makes a low return, the strategy moves like Grim

to a state with no investment and stays in this state for the remainder of the game. The strategies that best modeled participants' decisions in the role of player B simply made a constant return of either 50 or 67%. Thus overall these experimental results partly matched our simulation results, supporting Min-Grim and (nearly) Min-Return as strategies that can be—and often are—used to achieve stable cooperation in the IIG.

### 6.5. *Conclusions*

Recently several authors have argued that the decision strategies people use depend on the social domains they encounter (Bugental, 2000; Fiske, 1992; Hirschfeld and Gelman, 1994). Bugental (2000), for instance, proposes a distinction between five domains including hierarchical and reciprocity-based relationship. One important dimension separating domains is the structures of decision consequences or payoffs. A reciprocity-based domain might be characterized by a symmetric payoff structure, whereas a hierarchical power domain might be characterized by an asymmetric payoff structure. Following this reasoning, one could argue that the rules used in each domain could be shaped (e.g. through cultural evolution or learning) by the different decision structures that apply in each. In a reciprocity domain with symmetric decision consequences, a TFT-like strategy could be expected to emerge. In contrast, in a hierarchical power domain with asymmetric decision consequences like the IIG situation, a Grim-like strategy that is less cooperative is more likely to appear. However, these conclusions should be taken with some caution, since we have only examined one particular asymmetric game and some of its variants. This leaves open the possibility that in other asymmetric interactions strategies could emerge that are as cooperative as their counterparts in symmetric interactions. Nevertheless, in moving attention away from the often-studied symmetric interactions embodied in the IPD to asymmetric interactions, captured in the IIG, we have found a different set of strategies that may underlie many of the social interactions in which humans engage. These strategies

might often be not as “nice” as those that have been studied for symmetric interactions: even when efficient high-payoff outcomes are produced, the division of resources is unfairly skewed in one agent’s favor, reflecting the original asymmetry that the two parties brought to the interaction.

Furthermore, the efficient (if unequal) cooperative outcomes in an asymmetric situation such as the IIG are fragile. Once cooperation breaks down through an instance of exploitation, there is no good way for it to return. Player B cannot be contrite, because there is no way to signal to player A the intention to return to cooperative play if A does not first trustingly invest. Yet why should A do so and risk being exploited further if there is no evidence of contrition first forthcoming from B? Imagine the employee who has frequently worked overtime, trusting that his effort will be reciprocated with a bonus from his employer. After never being rewarded for this extra investment, should the employee spend any additional effort again at the next opportunity? Our results lead us to predict that cooperation can emerge in symmetric and asymmetric social interactions, but that it is more vulnerable to collapse in asymmetric interactions. If only one individual has the possibility of exploiting the other individual, it can be predicted that exploitation becomes more destructive. And when efficient outcomes do appear, one player is likely to obtain most of the surplus. Thus, despite the picture of even-handed and honorable cooperation that often emerges from studies using the symmetric Prisoner’s dilemma, the situation that arises in commonly-occurring asymmetric relationships may be as cooperative, but less equitable.

#### ACKNOWLEDGEMENTS

We gratefully acknowledge helpful comments on previous versions of this article by Carola Faselow, Scott Fisher, Richard McElreath, Jason Nobel, Leonhard Prechtel, and Anita F. Todd. Correspondence concerning this article should be addressed to Jörg Rieskamp.

## NOTES

1. Here we restrict our discussion to pure strategies—meaning that players use one strategy exclusively. However, the definition also holds for mixed strategies. A mixed strategy is a probability distribution over the set of pure strategies, which defines the probability with which each of the pure strategies is selected by the player.
2. A modification of the aspiration level has a large impact on the automaton's behavior, as it determines, for all states, to which state the automaton moves next depending on the opponent's decision. Because of this strong effect, the mutation function used a normal distribution with the old aspiration as the mean, thereby only changing the aspiration levels moderately.
3. In addition, we ran simulations for longer periods, but this did not lead to substantially different results. For instance, we ran one single evolutionary process for the IIG for 100,000 generations and determined for each 100th generation whether the strategies of the population produced an efficient outcome, that is a mutual payoff of at least 38. In 28.7% of these 1000 sampled generations the outcome was efficient, which is similar to the proportion of efficient outcomes when comparing the 100 independent runs with 1000 generations that we focus upon in the main text. Likewise we ran one single evolutionary process for the IPD for 100,000 generations and determined for each 100th generation whether the strategies of the population produced an efficient outcome (at least 38). In none of these generations were efficient outcomes observed, which again is similar to results obtained for the last 1000th (final) generation of the 100 independent runs we focus upon in the main text.
4. Here we assume that the number of states of the minimized automaton can be used as an indicator of an automaton's complexity: the higher the number of states, the more information an automaton can store, thus providing a larger memory with which to instantiate more complex strategies (see also Abreu and Rubinstein, 1988).

## REFERENCES

- Abreu, D. and Rubinstein, A. (1988), The structure of Nash equilibrium in repeated games with finite automata, *Econometrica* 56, 1259–1281.
- Aktipis, C. A. (2004), When to walk away: contingent movement and the evolution of cooperation, *Journal of Theoretical Biology* 231, 249–260.
- Alexander, R. D. (1987), *The Biology of Moral Systems*, Aldine de Gruyter, New York.



- Aumann, R. J. (1981), Survey of repeated games, In *Essays in Game Theory and Mathematical Economic in Honor of Oskar Morgenstern*, Bibliographisches Institut. Wissenschaftsverlag, Mannheim, pp. 11–42.
- Axelrod, R. (1984), *The Evolution of Cooperation*, Basic Books, New York.
- Axelrod, R. (1987). The evolution of strategies in the iterated Prisoner's dilemma, in Davis, L. (ed.), *Genetic Algorithms: A Simulated Annealing*, Pitman, London, pp. 32–41.
- Bendor, J. and Swistak, P. (1998), Evolutionary equilibria: characterization theorems and their implications, *Theory and Decision* 45, 99–159.
- Berg, J., Dickhaut, J. W. and McCabe, K. A. (1995), Trust, reciprocity, and social history, *Games and Economic Behavior* 10, 122–142.
- Binmore, K. (1994), *Game Theory and the Social Contract. Volume 1 Playing fair*, MIT Press, Cambridge, MA.
- Binmore, K. and Samuelson, L. (1992), Evolutionary stability in repeated games played by finite automata, *Journal of Economic Theory* 57, 278–305.
- Boyd, R. (1989), Mistakes allow evolutionary stability in the repeated Prisoner's dilemma game, *Journal of Theoretical Biology* 136, 47–56.
- Boyd, R. and Lorberbaum, J. P. (1987), No pure strategy is evolutionarily stable in the repeated Prisoner's dilemma game, *Nature* 327, 58–59.
- Bugental, D. B. (2000), Acquisition of the algorithms of social life: a domain-based approach, *Psychological Bulletin* 126, 187–219.
- Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- Deutsch, M. (1975), Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues* 31, 137–149.
- Fehr, E., Gächter, S. and Kirchsteiger, G. (1997), Reciprocity as a contract enforcement device: Experimental evidence, *Econometrica* 65, 833–860.
- Fiske, A. P. (1992), The four elementary forms of sociality: framework for a unified theory of social relations, *Psychological Review* 99, 689–723.
- Fudenberg, D. and Tirole, J. (1991), *Game Theory*, MIT Press, Cambridge, MA.
- Goldberg, D. E. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA.
- Goldberg, D. E. and Deb, K. (1991), A comparison of selection schemes used in genetic algorithms. In Rawlins G. J. E. (ed.), *Foundations of Genetic Algorithms*, Morgan Kaufmann, San Mateo, CA, pp. 69–93.
- Güth, W. and Kliemt, H. (2000), Evolutionarily stable co-operative commitments, *Theory and Decision* 49, 197–221.
- Güth, W., Ockenfels, P. and Wendel, M. (1997), Cooperation based on trust. An experimental investigation, *Journal of Economic Psychology* 18, 15–43.

- Güth, W., Schmittberger, R. and Schwarz, R. (1982). An experimental analysis of ultimatum bargaining, *Journal of Economic Behavior and Organization* 3, 367–388.
- Hirschfeld, L. A. and Gelman, S. A. (eds.) (1994), *Mapping the Mind: Domain Specificity in Cognition and Culture*, Cambridge University Press, New York.
- Hoffmann, R. (1999), The independent localisations of interaction and learning in the repeated Prisoner's Dilemma, *Theory and Decision* 47, 57–72.
- Hoffmann, R. (2001), The ecology of cooperation, *Theory and Decision* 50, 101–118.
- Hopcroft, J. E. and Ullman, J. D. (1979), *Introduction to Automata Theory, Languages, and Computation*, Reading, MA: Addison-Wesley.
- Kraines, D. and Kraines, V. (1993), Learning to cooperate with Pavlov: an adaptive strategy for the iterated Prisoner's Dilemma with noise, *Theory and Decision* 35, 107–150.
- Kraines, D. and Kraines, V. (1995), Evolution of learning among Pavlov strategies in a competitive environment with noise, *Journal of Conflict Resolution* 39, 439–466.
- Ledyard, J. O. (1995), Public goods: a survey of experimental results, in Kagel, J. H. and Roth, A. E. (eds.), *The Handbook of Experimental Economics*, Princeton University Press, Princeton, NJ, pp. 111–194.
- Leimar, O. (1997), Repeated games: a state space approach, *Journal of Theoretical Biology* 184, 471–498.
- Lindgren, K. (1991), Evolutionary phenomena in simple dynamics, in Langton, C. G., Taylor, C., Farmer, J. D. and Rasmussen, S. (eds.), *Artificial Life II*, Addison-Wesley, Reading, MA, pp. 295–312.
- Linster, B. G. (1992), Evolutionary stability in the infinitely repeated Prisoner's dilemma played by two-state Moore machines, *Southern Economic Journal* 58, 880–903.
- Lorberbaum, J. (1994), No strategy is evolutionarily stable in the repeated Prisoner's dilemma, *Journal of Theoretical Biology* 168, 117–130.
- Luce, R. D., and Raiffa, H. (1957), *Games and Decision*, New York: Wiley.
- Marinoff, L. (1990), The inapplicability of evolutionarily stable strategy to the Prisoner's dilemma, *British Journal of Philosophy of Science* 41, 461–472.
- Maynard Smith, J. (1984), Game theory and the evolution of behaviour, *Behavioral and Brain Sciences* 7, 95–125.
- Maynard Smith, J. and Price, G. R. (1973), The logic of animal conflict, *Nature* 246, 15–18.
- Michalewicz, Z. (1996), *Genetic Algorithms + Data Structures = Evolution Programs (3rd edn.)*, Springer, Berlin.
- Mitchell, M. (1996), *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA.

- Nelson, R. J. (1975), Behaviorism, finite automata, and stimulus theory, *Theory and Decision* 6, 249–267.
- Nowak, M. A. and May, R. M. (1992), Evolutionary games and spatial chaos, *Nature* 359, 826–829.
- Nowak, M. A. and Sigmund, K. (1993), A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game, *Nature* 364, 56–58.
- Nowak, M. A. and Sigmund, K. (1994), The alternating Prisoner's Dilemma, *Journal of Theoretical Biology* 168, 219–326.
- Ochs, J. and Roth, A. (1989), An experimental study of sequential bargaining, *American Economic Review* 79, 355–384.
- Rieskamp, J. and Gigerenzer, G. (2005), *Heuristics for social interaction: How to generate trust and fairness*. Manuscript submitted for publication.
- Samuelson, L. (1991), Limit evolutionarily stable strategies in two-player, normal form games, *Games and Economic Behavior* 3, 110–128.
- Samuelson, L. (1997), *Evolutionary Games and Equilibrium Selection*, MIT Press, Cambridge, MA.
- Schelling, T. C. (1960), *The Strategy of Conflict*, Harvard University Press, Cambridge, MA.
- Selten, R. (1983), Evolutionary stability in extensive two-person games, *Mathematical Social Sciences* 5, 269–363.
- Selten, R. (1988), Evolutionary stability in extensive two-person games—Correction and further development, *Mathematical Social Sciences* 16, 223–266.
- Sober, E. (1992), Stable cooperation in iterated Prisoner's Dilemmas, *Economics and Philosophy* 8, 127–139.
- Sugden, R. (1986), *The Economics of Rights, Cooperation, and Welfare*, Basil Blackwell, Oxford.
- Suppes, P. (1969), Stimulus-response theory of finite automata, *Journal of Mathematical Psychology* 6, 327–355.
- Trivers, R. L. (1971), The evolution of reciprocal altruism, *The Quarterly Review of Biology* 46, 35–57.
- Van Huyck, J. B., Battalio, R. C. and Walters, M. F. (1995), Commitment versus discretion in the peasant-dictator game. *Games and Economic Behavior* 10, 143–170.
- Weibull, J. W. (1995), *Evolutionary Game Theory*, MIT Press, Cambridge, MA.
- Young, H. P. and Foster, D. (1991), Cooperation in the short and in the long run, *Games and Economic Behavior* 3, 145–156.

*Address for correspondence:* Jörg Rieskamp, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. Tel.: +49-30-82406214; Fax: +49-30-8249939; E-mail: rieskamp@mpib-berlin.mpg.de

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.