Chapter 10 Simple Strategies for Social Interactions

The first two experimental studies have shown how people trust each other and how they are motivated by fairness in one-shot and ongoing interactions. Although the studies could show that the modal distribution of payoff was consistent with the equity principles, the large variance of outcomes could not be explained with simple distribution principles. What is needed, then, are models that describe the dynamic decision process of repeated interactions. Therefore *study 4* has the goal to develop models of people's decision strategies in the indefinitely repeated investment game. These should answer how people react if the decisions fall short of their expectations.

10.1 Competing Models

In principle, the repeated game strategies in an indefinitely repeated game are infinite. The strategies can be quite complex and can specify how the individual reacts for every possible decision history. In the present study participants in the role of player A in the investment game can only send an integer percentage of their endowment and participants in the role of player B can only return any integer percentage of the trebled investment, so that each stage game has 10,101 possible outcomes[14] giving, for instance, 1.1×10^{40} outcomes for 10 periods, which demonstrates how complex strategies can be, if they wish to specify behavior for all possible eventualities. However, it is reasonable to assume that people do not specify different behavior for every eventuality but follow simple strategies. In the present study the fit of three competing models describing the behavior is compared. First, I explore the fit of *simple strategies*, second, a *learning model* is used to predict decisions, and third, a *baseline model* is used as a benchmark.

Simple Strategies for Social Interactions

Given the limits on human memory and computational capacities it appears reasonable to assume that individuals prefer simple strategies for decision making. Simple strategies may incorporate two advantages: First, the cognitive demands for applying them may be rather low, and second, they may be quite effective. Gigerenzer et al. (1999) showed that heuristics can be quite powerful in solving various judgment and decision

[14] Player A chooses from among 101 percentages how much he wishes to invest (0-100%). Player B can only make a decision if player A makes an investment, hence, for the 100 possible investments player B chooses from among 101 possible returns (0-101%). This gives 1+(100x101)=10,101 possible combinations.

problems. Other studies, most of them based on the use of the prisoner's dilemma, also demonstrated the good performance of simple strategies for social interaction.

The simulation studies reported in chapter 9 could show how simple strategies, like the Grim strategy, are sufficient to compete against other potentially more complex strategies for the indefinitely repeated investment game. Finite automata were used to represent repeated game strategies for the investment game. The use of finite automata as strategies for games in evolutionary simulations is an established practice. Using automata as models for individuals' decision processes, however, has rarely been done (for exceptions see Engle-Warnick & Slonim, 2001, and Wedekind & Milinski, 1996). In the present study I will propose automata for representing people's decision strategies in the investment game. In chapter 11 the strategies selected to predict individuals' decisions will be related to the ones found in the evolutionary simulations.

The automata that I used as models for individual's strategies in the investment game are the same as those used in the evolutionary simulations reported in chapter 9. Each automaton is restricted to a maximum number of 4 states. For simplicity an automaton's output is restricted to multiples of 10 ranging from 0 to 100 determining the investment or return rate. As these automata are used as models for participants (i.e. replacing participants) the input of the automata is given by the decisions of the participant in the other player role. Again based on various research (Simon, 1956; 1990; Lopes, 1996) claiming that individuals use aspiration levels for categorizing the possible options or consequences of a decision, the strategies use an aspiration level for categorizing investments or returns into two categories labeled as "Trust" or "Distrust" and "Reciprocity" or "Exploitation." Given this categorized input the automaton advances to another state, which determines the output (i.e. the decision in the next period). The decisions of the automata predict participants' decisions.

Belief-Based Reinforcement Learning Model

As a competing model the predictive strength of a learning model is explored. Learning models have recently attracted increasing interests in judgment and decision-making research (Busemeyer & Myung, 1992; Camerer & Ho, 1998, 1999a, 1999b; Erev & Roth, 1998; Fudenberg & Levine, 1998). There is an ongoing debate regarding the central assumption about the cognitive process these models should incorporate. One approach—belief-based models—assumes that individuals form expectations about the behavior of others involved in an interaction based on their past behavior. Depending on their expectation, they will then make the choice that maximizes their expected payoffs. By acquiring experience with the decision situation, the expectations are updated and consequently decisions are changed. A standard strategy of this approach is *fictitious play* (Brown, 1951), a strategy that simply selects the option that would have resulted in the highest payoff given the opponent's past behavior (see also Fudenberg & Levine, 1998). Another approach—reinforcement models—assumes that the probability that a particular option is selected by an individual depends on the reinforcement this option has obtained in the past (Erev & Roth, 1998; Roth & Erev, 1995). In the present study I will use a modified version of the "experience–weighted attraction model" proposed by Camerer and Ho (1999a), which I label the "experience–weighted expectancy model" (EWE). The fit of the EWE will be compared to the fit of the simple strategies. The model assumes that individuals have expectations about their opponents' behavior and based on their expectations each possible option is assigned a subjective value—expectancy. Initially, individuals will choose the option with the highest expectancy. Depending on the positive or negative consequences of their choices—the reinforcement—the expectancies of the options are updated during the decision process and consequently influence and change choices.

More formally, the learning model considered in the present study assigns expectancy (*E*) to each option. The model also assumes that individuals have an aspiration level with which the obtained payoff in each period is compared. For player A, the reinforcement *R* for a chosen option is given by the obtained payoff minus the aspiration level. For instance, if an individual has an aspiration level of 20 but obtains only a payoff of 10 this leads to a negative reinforcement of minus 10. Because player B makes a decision after player A's decision the payoff player B will obtain in the period is obvious. It is always apparent to player B that she will obtain a higher payoff if she does not return anything; however it is not clear whether player A will repeat a high investment in the subsequent period. Therefore, the success or reinforcement of player B's decision is evaluated by the amount of investment that is invested by player A in the next period. All options that were not selected are assigned a reinforcement of R=0. Here I differ from Camerer and Ho's (1999a) model that uses hypothetical payoffs, that is, payoffs that would have resulted from choosing other options assuming an unchanged decision of the opponent. Using hypothetical payoff for the investment game is problematic as players make decisions sequentially. Player B always takes the decision of player A into account; therefore, it is unjustified to assume that if player A had chosen another option, player B would not have changed her decision. The EWE also keeps track of how much experience has been acquired expressed in $N(t)$. Due to past experience and imagination, the initial experience does not have to start with a value of zero. Experience is updated after each period, according to Rule 1. A depreciation rate r measures the impact of past experience compared to new experience and is used as a weight for $N(t)$. After updating the experience, the expectancies are updated according to Rule 2. Here a discount factor f is used to depreciate previous expectancies (Roth and Erev, 1995, call this parameter "forgetting"). Because of potentially negative reinforcements the values for the expectancies could also be negative. In this case I define that negative expectancies are always replaced by an expectancy of zero.

Rule 1: Experience updating

$$N(t) = r \cdot N(t-1) + 1,\ \ t \geq 1$$

Rule 2: Expectancy updating

$$E_{t+1}(option) = \frac{f \cdot N(t-1) \cdot E_t(option) + R(option)}{N(t)}$$

Expectancies are usually used to determine choice probabilities. There are several ways that the expectancies could be transformed to probabilities such that the probability of choosing an option monotonically increases with the expectancies. Different transformations like exponential, normal, or power transformations are discussed in the literature (Camerer & Ho, 1999a; Erev & Roth, 1998; Fudenberg & Levine, 1998). However, for the present study, I will omit determining choice probabilities. To compare the fit of the learning model with the deterministic strategies, it is necessary to make deterministic predictions also for the learning model. Therefore, the option with the highest expectancy is predicted by the model.

Baseline Model

In addition to the simple strategies and the EWE model, a baseline model is used as a standard for comparison. The baseline model simply assumes that individuals select investment and return rates with a constant probability. Unlike the other models the baseline model predicts that the choices are independently and identically distributed across all games and periods. The baseline model is a strong competitor for the other models, because participants' decisions are used to fit the choice probabilities. The other models can only outperform the baseline model in predicting participants' decisions if they succeed in explaining how the decisions depend on the decisions of the opponent, and how decisions change from period to period.

In sum, in the present study, the fits of different models describing individuals' decision processes are compared. As these models and their parameters are fitted to the data they face the problem of "capitalization on chance" or "overfitting," that is, developing a model that fits the data but can hardly predict new data (see Browne, 2000). Therefore a cross-validation method is applied. First a calibration sample is used for the development of the different models and for estimating models parameters. These developed models are then used to predict participants' behavior in a validation sample. For assessing and comparing the validity of the models for describing individuals strategies, I suggest considering two aspects: first, the difference between models' fits in describing the data of the participants of the calibration sample, and second and more importantly, the fit in predicting the independent data of the validation sample, which provides the possibility of assessing the feasibility of generalizing the models.

In the present study participants again had to play the indefinitely repeated investment game. In contrast to study 2 reported in chapter 8, the endowment given to player B was not varied. Instead, all participants in the role of player B were provided with an endowment in each period. Another difference to study 2 was that participants had to play the investment game in both player roles. Thereby, the experimental procedure is consistent with the analytic exploration and the evolutionary simulations of study 3 in which the agents also had to play in both player roles. It is also explored whether this modification has an effect on the decision behavior. It may be possible that individuals show higher trust, reciprocity, and motivation for fairness if they act in both player roles, because they experience the perspective of their interaction partner. Following this argument, one would expect higher investment and return rates in the present study (for the effect of perspective taking in decision making see also Drolet, Larrick, & Morris, 1998; Gigerenzer & Hug, 1992; McPherson Frantz & Janoff-Bulman, 2000).

10.2 Method

In the present study, 24 men and 24 women, mainly students of the Free University of Berlin, participated in 12 sessions, each with 4 participants. The participants were not acquainted with one another. Players A and B were placed in various rooms and all interactions were performed via personal computers. Each participant played the investment game against all three other participants of the session, one at a time, in both player roles, resulting in six repeated games for each person. The randomly generated game lengths (with a continuation probability of 0.875) used for study 2 were also taken for the present study, yielding the same lengths for each session and comparability between the experiments. In each game and in each period both participants began with an endowment of DM 10.

First the investment game was explained to the participants. They were told that after each period with a probability of 0.875 a new period would follow. They were told that they had to play three games in both player roles against three other participants, whose identity would not be revealed, to guarantee anonymity. It was also explained to the participants that in each period of every game both players would start with an endowment of DM 10. Player A could invest any amount of his endowment and player B could (possibly) return any amount of the trebled investment. After the players had made their choices both players were informed about the decisions of the opponent and the corresponding payoffs gained by both players of the present period were displayed. Participants were not informed about their total gains accumulated over the last periods and games. Participants were told that two percent of the final payoffs would be given as payment. This resulted in a maximum payment of around real DM 40 (ca. $20), for a session duration of approximately one hour. The instructions used for the experiment were quite similar to the instructions used for study 2, which are presented in appendix B.

10.3 Results

Decisions

In 47% of all 1,152 periods the participants in the role of player A invested 100% of their endowment, thereby, producing efficient outcomes, that is, maximum mutual payoffs. On average participants in the role of player A accumulated DM 14.9 across all repeated games, whereas participants in the role of player B accumulated DM 19.3. The modal outcome was an efficient equal payoff of DM 20 for both players (8% of all 144 repeated games).

Figure 15 shows the frequency distribution of outcomes. The most frequent outcomes were an equal split of the final payoffs (20, 20; with a frequency of 19%), an equal split of the trebled investment (15, 25; with a frequency of 14%), and final payoffs without any surplus (10, 10; with a frequency of 12%). All other outcomes occurred with frequencies lower than 4%. In the cases where the entire endowment was invested by player A, equal final payoffs (20, 20) were obtained substantially more frequently than equal splits of the trebled investment (15, 25), consistent with the equity prediction that says individuals strive for final equal payoffs.

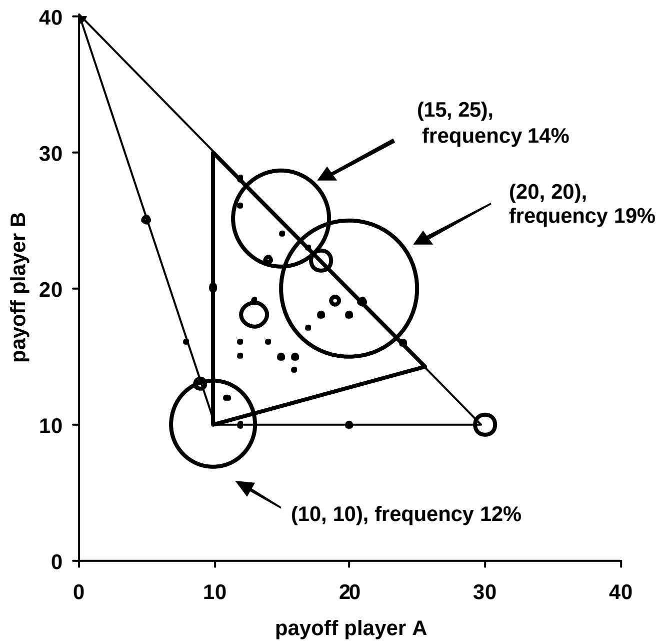

Figure 15. The graph shows the distribution of payoff combinations across all 1,152 periods of the experiment. The centers of the circles represent the outcomes that are reached by the participants. Circles' diameters are proportional to the frequency with which the outcomes occurred across all periods. (The triangle formed by the bold lines indicates payoff pairs predicted by the folk theorem of indefinitely repeated games.)

To test whether the role change included in the present study had any effect on participants' decisions the average investment and return rates of the sessions were compared with those in study 2 for the Endowment condition. No substantial differences between the investment rates and return rates could be observed. The average investment rates of 71% (*SD*=15) in study 2 and in the present study with 71% (*SD*=13) did not differ. The average return rate of 54% (*SD*=9) in study 2 tended to be higher than in the present study with an average return rate of 48% (*SD*=13), (t_{16}=1.1, *p*=.300; medium effect size *d*=0.54).

Again I tested prediction 7 whether investment rates are correlated with return rates. For determining the correlation only one game per person was used to provide independent observations and all periods in which no investment was made were omitted (since no return could be made either, which would artificially increase the correlation). Because participants played the game in both player roles, for each participant either the decision for the role of player A or for player B was selected, yielding altogether 24 independent observations. Consistent with prediction 7.2 a substantial correlation between the average investment rate and the average return rate of r=.48 was found (p=.017; large effect size). Thereby the present study could again demonstrate positive and negative reciprocity in ongoing interactions. Given that the correlation is larger in the present study compared to study 2 one can speculate that the role change conducted in the present study supports reciprocal behavior.

To acquire a representation of the dynamics of the games, participants' investments and returns in the first and fourth periods of the game were analyzed (the fourth period was the highest number of periods all games were repeated). In the first period the average investment rate was 74% compared to an average investment rate of 72% in the fourth period. In the first period the average return was 54% compared to an average return of 45% in the fourth period. This illustrates that contrary to the investment rates, the return rates declined during the game.

Decision Models

Three models were developed to describe participants decisions. As these models and their parameters were fitted to the data they face the problem of "capitalization on chance." Therefore a cross-validation method was applied, so that a calibration sample of eight randomly selected sessions with 32 participants was used for developing models. Subsequently the models were used to predict the data of the remaining four sessions with 16 participants, which posed as a validation sample. To evaluate the models the differences in their fits for the calibration sample and, more importantly, their predictions for the validation sample are considered (for an easier comparison of the models their fits are summarized in Table 5 on page 98).

Baseline model.

The baseline model assumes that individuals have constant probabilities with which they select the possible alternative options. For simplification the possible predictions of all models are restricted to multiples of 10 with a minimum of 0 and a maximum of 100 (which yields 11 possible predictions for each period). The fit of all models including the baseline model is defined as the percentage of participants' decisions that can be correctly predicted (with a tolerated deviation of ±5%). When evaluating the models' fit note that a model that simply randomly selects 1 of the possible 11 predictions with equal probability would by chance predict on average about 9% of the decisions.

The baseline model was fitted to the data of the calibration sample and predicts the most frequent decision of the participants in the role of player A (player B). The model for player A predicts an investment of 100% in all periods of all games and reaches a fit of 52% for the calibration sample. For the purpose of cross-validation the baseline model predicted 44% of participants' decisions in the role of player A in the validation sample. The model for player B predicts a return of 50% in all periods, thereby reaching a fit of 34% for all decisions in the calibration sample. The baseline model for player B predicted 29% of the decisions in the validation sample.

Simple strategies.

Finite state automata were developed to describe people's decision strategies. To build adequate automata as decision strategies a genetic algorithm was applied (Goldberg, 1989; Mitchell, 1996). In contrast to the genetic algorithms used in chapter 9, in which strategies compete against each other, in the present study the strategies' goal was to predict participants' decisions as accurately as possible. Studies by Wedekind and Milinski (1996) or Engle-Warnick and Slonim (2001) reported in chapter 4 demonstrate that different individuals use different strategies. Therefore, the genetic algorithm always produced a set of two strategies. The fit of each strategy is defined as the percentage of participants' decisions that can be correctly predicted (with a tolerated error of ±5%). Each participant is assigned the strategy with the highest fit out of the set of two strategies. The fit of the set of two strategies is defined as the average fit the strategies reached that were assigned to the participants. For example, it could be the case that one strategy is used as a decision model for 20 participants and the other as a model for 12 participants. The average fit of the set of strategies is then determined across the fit of the first strategy for the 20 participants and the fit of the second strategy for the other 12 participants.

The genetic algorithm started at the outset to produce a population of pairs of strategies, by choosing the parameters values of the strategies randomly with equal probability. Then the fit of all pairs of strategies in predicting participants' decisions was determined. Subsequently a new population of pairs of strategies was selected for the next generation. Those pairs of strategies that reached a higher fit were more likely to be selected for the next generation. Eventually automata are modified by a *crossover* and/or a *mutation* procedure and thereby potentially improve in predicting participants' decisions. For each player role five genetic algorithms were run in which the developed strategies replaced participants in the role of either player A or player B. The pair of strategies with the highest average fit of the five runs in the last generation was selected as the model for participants' decision strategies (for details on the genetic algorithm see appendix C).

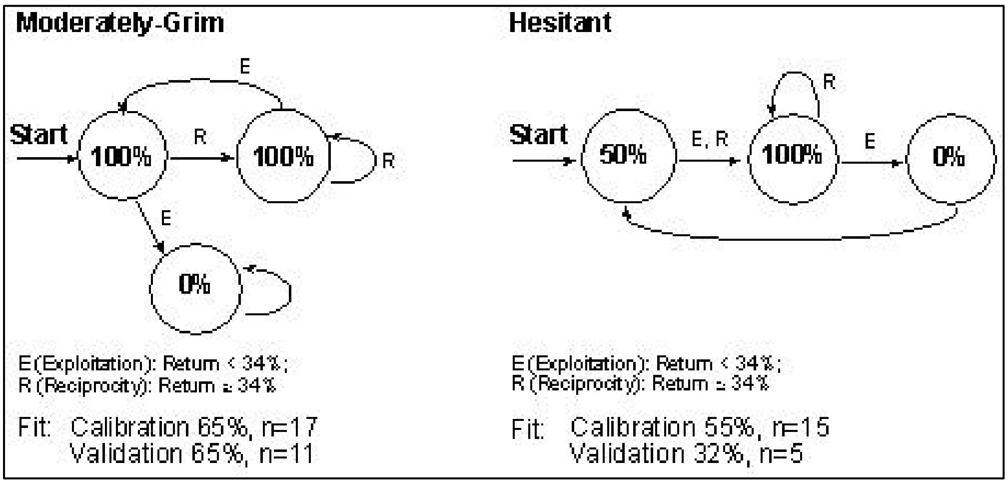

Figure 16. The pair of strategies for player A with the highest average fit of 61% for the calibration sample, which could predict 55% of all decisions of participants in the role of player A in the validation sample.

Figure 16 shows the pair of strategies that reached the highest average fit of 61% in predicting participants' decisions in the role of player A for the calibration sample. For the validation sample the strategies could predict on average 55% of participants' decisions.

The first strategy, *Moderately-Grim*, consists of three states and could predict 65% of all decisions for 17 participants to whom it was assigned in the calibration sample. For the validation sample it reached a fit of 65% for 11 participants. Moderately-Grim commences with an investment of 100% and if player B returns an amount greater than or equal to 34% it moves to state 2, in which it repeats the investment of 100% as long as player B makes a substantial return. However, if in the first period player B makes a return lower than 34% Moderately-Grim advances to the third state, in which it will make no investment in the next and in all following periods. If the opponent returns less than 34% two times consecutively, the strategy always ends in the third state, which is a terminal state as it is not possible to move back to any other state.

The second strategy, *Hesitant*, consists of three states and reached a fit of 55% for 15 participants in the calibration sample. Hesitant could predict 32% of 5 participants' decisions in the validation sample. Hesitant starts in the first period with a moderate investment of 50% and increases its investment in the second period to 100%. It remains in the second state with an investment of 100% as long as the opponent returns at least 34%

of the trebled investment. If the opponent returns less, then Hesitant advances to state 3, where no investment is made, and subsequently it reverts to state 1.

Figure 17 shows the pair of strategies that reached the highest average fit of 58% in predicting participants' decisions in the role of player B for the calibration sample. For the validation sample the strategies could predict on average 50% of participants' decisions.

The first strategy, *Reactive*, consists of two states and could predict 59% of all decisions for 16 participants to whom it was assigned in the calibration sample. For the validation sample it was assigned to 7 participants and could predict 52% of their decisions. Depending on player A's decision in the first period, Reactive either returns nothing if the investment is lower than 17% or returns 70% given a higher investment. Reactive will always move to the first state with no return if player A makes a low investment and it always moves to state 2 with a return of 70% if the player A invests at least 17% of the endowment.

The second strategy, *Half-Back*, consists of two states and could predict 57% of all decisions for 16 participants to whom it was assigned in the calibration sample. For the validation sample it predicts 47% of the decisions of 9 participants. Regardless of player A's first decision, Half-Back returns 50% of the trebled investment in the first period and repeatedly returns 50% in all subsequent periods, if player A repeatedly invests more than 12% of the endowment. If the investment rate falls below 12%, Half-Back returns nothing but will always revert to a 50% return rate in the following period.

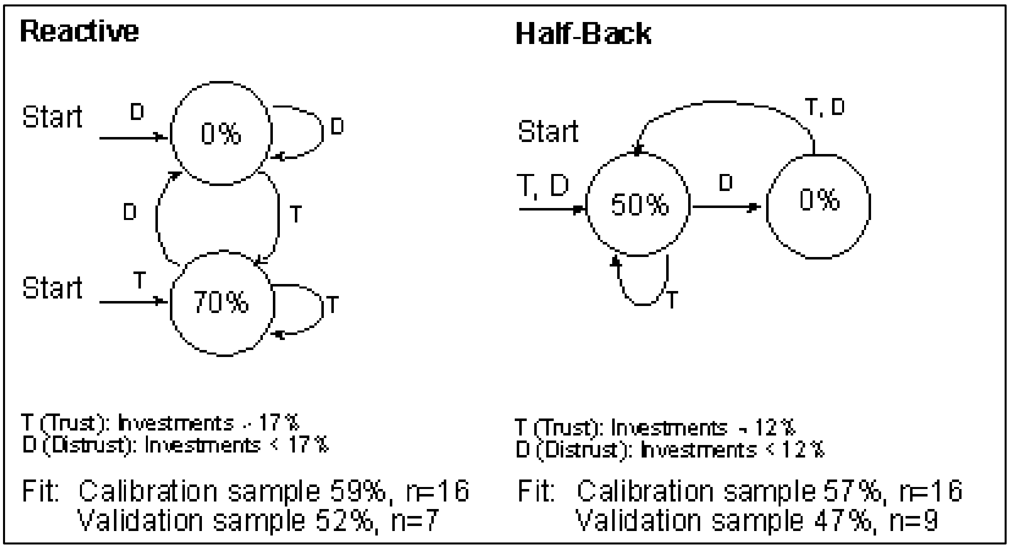

Figure 17. The pair of strategies for player B with the highest average fit of 58% for the calibration sample, which predicts 50% of all participants' decision in the role of player B in the validation sample.

Experience-weighted expectancy learning model.

Due to the large number of parameters of the EWE model again a genetic algorithm was used to find the best parameters. For simplicity only investment and return rates divisible by 10 were used as options and expectancies for these 11 options were determined. The expectancies (E) were restricted to the maximum payoffs reachable for the players, that is, for player A $E \in [0,30]$ and for player B $E \in [10,40]$. These intervals were chosen to connect expectancies to the possible payoffs that could be obtained by player A and player B. The aspiration level was restricted to integers ranging from 0 to 30 for both players. Note for the EWE model for player A the level refers to the amount returned by player B and for the EWE model for player B the level refers to the trebled amount of investment sent by player A in the next period. The other parameters were restricted as follows: depreciation rate $r \in [0,100]$, initial experience $N(0) \in [0,(1/(1-r)]$, and discount factor $f \in [0,1]$.

At the outset the genetic algorithm used randomly chosen values as parameter values for 60 models. The models are used to predict participants' decisions. After each period of the game models' parameters are updated by using the real decisions of the participants. For each participant the percentage of correctly predicted decisions was determined and the average fit across all participants was computed. Models with a higher average fit were more likely to be selected for the next generation of models. Therefore during a developmental process, in which the parameter values were changed by a mutation and crossover procedure, the parameter values were optimized. To avoid parameter values of local maxima the genetic algorithm was run repeatedly and the EWE model with the highest fit was finally selected.

The EWE model for player A reached a fit of 58% in predicting participants' decisions in the calibration sample. For the validation sample the EWE model could predict 51% of all decisions. Table 3 shows the values for the initial expectancies. Investments of 100% reached the highest expectancies, but the expectancy for the 0% option was nearly as high. The other parameter values were aspiration level $a=8$, depreciation rate $r=1.18$, initial experience $N(0)=1.71$, and discount factor $f=0.03$.

Table 3

Expectancies of the EWE model for the different investment options of player A

	Options – Investment Rates										
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Expectancies	23.7	19.0	9.0	14.8	22.9	17.2	3.2	20.1	13.8	14.8	24.1

Note: Expectancies were restricted to values ranging from 0 to 30.

The EWE model predicts that in the first period an investment of 100% is made. In the second period the prediction depends on the reinforcement, that is, the return rate. If player B returns more than the specified aspiration level of 8 then this leads to a positive reinforcement of the selected investment rate. In contrast, a lower return than 8 produces a negative reinforcement, that is, the expectancy of the selected investment rate is reduced. The model always predicts an investment of 100% as long as player B makes a return above the aspiration level. In contrast, a return lower than the aspiration level in the first period reduces the expectancy for the 100% investment option, yielding an investment of 0% in all following periods. Which investment rate is chosen depends on the expectancies of the investment rates. If, for instance, in the first period player B makes a high return of, say, 70%, then this increases the expectancy for the 100% investment option so that even if player B returns nothing in the following periods the EWE will still predict an investment of 100% in the following periods. Only if player B subsequently makes no returns in three following periods, the EWE model predicts that no investment will be made in the following period.

The EWE model for player B reached a fit of 47% in predicting the decisions of the participants in the calibration sample. For the validation sample the EWE model predicts 39% of participants' decisions. Table 4 shows the values for the initial expectancies. A return rate of 70% reached the highest expectancy. The other parameter values were aspiration level $a=0$, depreciation rate $r=14.3$, initial experience $N(0)=0.01$, and discount factor $f=0.74$.

Table 4

Expectancies of the EWE model of the different return options for player B

	Options – Investment Rates										
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Expectancies	36.0	31.4	34.5	30.9	26.4	28.5	26.7	36.2	24.0	16.4	21.9

Note: Expectancies were restricted to values ranging from 0 to 30.

The EWE model predicts that in the first period a return of 70% is made. Given the low aspiration level of 0 every possible investment leads to a positive reinforcement of the 70% return rate, so that the model constantly predicts a return rate of 70%.

In summary, three different models were developed by using the data of a calibration sample. The developed models were then used to predict the new, independent data, that is, the participants' decisions in a validation sample. For a better comparison the fits of the different models are summarized in Table 5. It can be seen that the simple strategies outperform the baseline model and the EWE learning model in predicting participants' behavior in both the calibration and the validation sample.

Table 5

Models' fit in predicting participants' decisions in the role of player A and player B in the repeated investment game, separated for the calibration and validation samples

	Player A		Player B	
Model	calibration	validation	calibration	validation
Baseline Model	52 %	44 %	34 %	29 %
Simple Strategies	61 %	55 %	58 %	50 %
EWE - Learning	58 %	51 %	47 %	39 %

10.4 Discussion

Stable fairness principles are not sufficient to explain peoples' decisions in a dynamic situation like the repeated investment game. Therefore the goal of the present study was to propose models of peoples' decision strategies. These models should define the conditions when a person trusts his or her opponent and on the other hand when a person reciprocates trust with a substantial return.

Similar to the results found in study 2, substantial trust and reciprocity were observed in the repeated game. The modal outcome across all periods was an efficient equal payoff of DM 20 for both players. In all cases in which player A invested his entire endowment an equal final payoff was obtained substantially more frequently than equal splits of the trebled investment (15, 25), consistent with the equity prediction.

One may speculate that the role changes conducted in the present study allowed the participants to take different perspectives, and that these perspective changes may have influenced participants' behavior. For instance, if one takes the role of player B and never returns anything to player A then this behavior may also lower the expectations of high returns when taking the role of player A. However, the investment rates did not differ compared to the investment rates of the corresponding condition in study 2. The average return rates also did not differ substantially between the two studies. In contrast, the correlation between the investment rates and the return rates was higher in the present study than in study 2. Therefore, it can be argued that the role changes support positive and negative reciprocity. By changing individuals' perspectives in the game individuals who first take the role of player A and make high investments might in subsequent games expect high investments from other participants; if these are not made they are punished with low returns. Participants who only take the role of player B might not have strong expectations about player A's behavior, thereby they do not punish low investments with low returns.

The main part of the present study was to investigate individuals' decision processes. To evaluate the different models of the decision processes their fit could be compared to each other, and additionally the predictive capacity was tested with a validation sample. Both the simple strategies and the learning model outperformed the baseline model in the calibration sample. However, in the validation sample only the simple strategies outperformed the baseline model. Given that the baseline model that only predicts one single decision for all decisions reached the same fit as the EWE model, this clearly disqualifies the EWE model for player A. For player B the EWE model outperformed the baseline model in predicting participants decisions, but the simple strategies could predict a substantially larger percentage of participants' decision compared to the EWE model. Therefore the simple strategies are the best models for describing a substantial percentage of individuals' decisions.

In additional to considering the fit of the two approaches one should consider the assumptions they make about the cognitive process and their plausibility. The learning model makes two basic assumptions: First, it assumes that for all possible options subjective values are formed, and second, that these expectations are updated based on experience or reinforcement. The EWE model assumes that the chosen options are updated according to the reinforcement and all other options are updated with a fixed value. Other authors propose that in addition, unselected options are updated by hypothetical reinforcement, which was not realized (Camerer & Ho, 1999a). Such an updating process requires the decision maker to think through all hypothetical cases and compute the forgone payoffs of the not-chosen options. The updating process would require a lot of computations in every period of the game, which appears less realistic, at least as a conscious cognitive process.

Besides the psychological plausibility of the updating process, there is a more severe problem for learning models that concerns the ability to model sudden changes in behavior caused by strategic considerations of the decision maker. Learning models have their strengths in a situation in which the optimal option out of a set of options has to be found, like in a probability learning task, or in the situation of a zero-sum game, in which individuals have no possibility to increase their payoffs by cooperation. However, in a situation in which the decision maker follows a certain strategy over several periods this can hardly be modeled with a learning model that focuses on the success of behavior in single periods. A vivid example is a strategy that alternates between two options. Such a strategy could be followed in a repeated chicken game (see Figure 4 on page 19) in which the alternation between options leads to an efficient equal payoff for both players. In such a situation not the single payoff in one period but the conjoint payoffs over two periods is decisive. Such an alternation between options could easily be represented with a two-state automaton. The automaton simply alternates between the two states: one state leads to the selection of the first option and the output of the other state leads to the selection of the other option.

Another example that is difficult to model with a learning model can be illustrated with the investment game. Someone in the role of player A might follow the strategy of beginning with a small investment and raising this investment in small steps if the investment is reciprocated with substantial returns. This strategy is difficult to model with a pure reinforcement strategy, because every choice leads to a positive reinforcement, which reinforces the selected option and restrains other options. A belief-based learning model that uses hypothetical payoffs would also run into problems, because it would predict a strategy change in favor of the option that would produce the highest possible payoff, whereas the decision maker prefers the option with a small increase in payoffs. These examples demonstrate the difficulties of modeling repeated game strategies with learning models.

Building Blocks of the Simple Strategies

Individuals who apply Moderately-Grim (or have a decision strategy that can be best described with it) are initially very confident in trusting their opponents and attempt to construct a mutually beneficial relationship. The strategy commences with *initial trust* by investing the entire endowment. It continues investing the entire endowment if the opponent returns a substantial proportion of the trebled investment. Even if individuals who apply such a strategy receives no return from their opponent they will still invest their entire endowment, thereby being quite *forgiving*. Only if they are exploited twice will they not invest anything in any subsequent period, thereby applying *unforgiving punishment*. If Moderately-Grim is analyzed game-theoretically there are "best reply strategies" that lead to a low payoff for Moderately-Grim. For instance, player B could return only 34% of the trebled investment in the first period and return nothing every other move.

The second strategy Hesitant found for player A is even more forgiving, because it has no *unconditional punishment* state. The strategy is cautious by making only a small investment in the first period and increases this investment in the following period. In the next period an investment of 100% is made, which is repeated in all following periods if player B returns at least 34% of the trebled investment. A subsequent low return is punished with no investment in the following period. However, thereafter the strategy always tries to build up cooperation again. A best reply strategy for Hesitant should, after the first period, make a return above Hesitant's aspiration level. Because the fit of the Moderately-Grim strategy does not decrease for the validation sample, contrary to that of Hesitant, it should be regarded as the more robust strategy for predicting individuals' decisions.

Participants' behavior in the role of player B was described with two strategies: Reactive and Half-Back. If player A makes large investments then Reactive makes a return that leads to an equal payoff for both players; thereby it explain when a fair outcome, according to the equity principle, is obtained. Low investments are punished with no returns at all. Half-Back returns 50% of the trebled investment given an investment of at least 12%, revealing a self-serving tendency as it leads to a higher payoff for player B compared to player A (for self-serving tendencies see also Kagel, Kim, & Moser, 1996). For both strategies a best reply strategy for player A is to invest always the entire endowment, as Moderately-Grim does. Together both strategies can explain the two most frequently observed outcomes (15, 25 and 20, 20) in the study. Beyond this they can explain under which conditions other outcomes are accomplished. In sum, when interpreting the strategies' mechanisms individuals appear to reciprocate kind behavior, but some do not hesitate to make some extra gain.