

Wright State University

CORE Scholar

Psychology Faculty Publications

Psychology

9-2015

Modeling Trust Dynamics in Strategic Interaction

Ion Juvina

Wright State University - Main Campus, ion.juvina@wright.edu

Christian Lebiere

Cleotilde Gonzalez

Follow this and additional works at: <https://corescholar.libraries.wright.edu/psychology>



Part of the [Psychology Commons](#)

Repository Citation

Juvina, I., Lebiere, C., & Gonzalez, C. (2015). Modeling Trust Dynamics in Strategic Interaction. *Journal of Applied Research in Memory and Cognition*, 4 (3), 197-211.

<https://corescholar.libraries.wright.edu/psychology/420>

This Article is brought to you for free and open access by the Psychology at CORE Scholar. It has been accepted for inclusion in Psychology Faculty Publications by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.



Original Article

Modeling trust dynamics in strategic interaction

Ion Juvina^{a,*}, Christian Lebiere^b, Cleotilde Gonzalez^c^a Department of Psychology, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435, USA^b Department of Psychology, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA^c Department of Social and Decision Sciences, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:

Received 17 October 2013

Accepted 30 September 2014

Available online 22 October 2014

Keywords:

Cognitive modeling

Trust dynamics

Social Dilemmas

Strategic interaction

Transfer of learning

ABSTRACT

We present a computational cognitive model that explains transfer of learning across two games of strategic interaction – Prisoner's Dilemma and Chicken. We summarize prior research showing that, when these games are played in sequence, the experience acquired in the first game influences the players' behavior in the second game. The same model accounts for human data in both games. The model explains transfer effects with the aid of a trust mechanism that determines how rewards change depending on the dynamics of the interaction between players. We conclude that factors pertaining to the game or the individual are insufficient to explain the whole range of transfer effects and factors pertaining to the interaction between players should be considered as well.

© 2014 Society for Applied Research in Memory and Cognition. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Humans have a strong tendency toward cognitive parsimony: they tend to develop cognitive strategies that make only minimal use of the potentially relevant information in the environment (Gigerenzer, Todd, & the ABC Research Group, 1999). In doing so, they do not compromise their ability to adapt and thrive, quite the contrary. For example, a trust heuristic assists us in dealing with the complexities of interpersonal interaction (e.g., Wegwarth & Gigerenzer, 2013). Once we have identified a trustworthy person, we tend to suspend the meticulous analysis of the benefits and risks of cooperating with that person; we just assume (i.e., trust) that he or she will reciprocate in kind. Applying a heuristic (i.e., simple rule) can speed up decision-making and reduce cognitive load, releasing cognitive resources that allow us to adapt to complex and dynamic environments. Forgoing meticulous analysis and relying on simple rules derived from experience is one of the characteristics of intuition (Gigerenzer, 2007). Intuitive decision making can be very effective in handling the complexity and uncertainty of social environments by exploiting evolved capacities and environmental regularities (Hertwig & Hoffrage, 2013). Here we use computational cognitive modeling to investigate how the coupling between simple heuristics, cognitive capacities, and social environments might work in strategic interpersonal interaction. We build

on previous research suggesting that cognitive architectures and particularly instance-based learning (IBL) approaches may provide a general explanation of intuitive decision-making (Gonzalez, Ben-Asher, Martin, & Dutt, 2015; Gonzalez, Lerch, & Lebiere, 2003; Thomson, Lebiere, Anderson & Staszewski, 2015).

Games of strategic interaction have successfully been used to model various real-world phenomena. For example, the game Prisoner's Dilemma has extensively been used as a model for real-world conflict and cooperation (Rapoport, Guyer, & Gordon, 1976). These games are often called *social dilemmas* to emphasize their relevance for the real world. There has been a recent tendency toward studying ensembles of games, as most social dilemmas rarely occur in isolation; more often they take place either concurrently or in sequence (Bednar, Chen, Xiao Liu, & Page, 2012). This is particularly true in organizations with complex structures, roles, and processes. For instance, when games are played in sequence (i.e., one after another), an effect known as “spillover of precedent” may occur: a precedent of efficient play in a game can be transferred to the next game (e.g., Knez & Camerer, 2000). We refer here to games that are repeated multiple times; the players acquire extensive experience with one game before they switch to another game. We determine the effect that the first game has on the second one and we refer to this effect as *transfer of learning in games of strategic interaction*. This effect has important practical implications. For example, most organizations employ training exercises to develop cooperation and trust among their employees. The assumption is that what is learned in a very specific, ad-hoc exercise transfers to organizational life once the training is over. Much of expertise is generally of an intuitive nature (Gigerenzer, 2007), which makes it

* Corresponding author. Tel.: +1 9377753519.

E-mail addresses: ion.juvina@wright.edu (I. Juvina), cl@cmu.edu (C. Lebiere), coty@cmu.edu (C. Gonzalez).

inaccessible to conscious thought and thus hard to study with traditional methods like self-reporting. Here, we employ a cognitive architectural approach to analyze the interplay of cognitive processes and interaction dynamics that underlie what appears as gut feelings or intuition.

Research in behavioral game theory attempting to explain what causes transfer of learning in games of strategic interaction can be summarized as follows: (1) [Bednar et al. \(2012\)](#) use the concept of *entropy* or strategic uncertainty to explain when learned behavior is likely to spillover from one game to another. They suggest that prevalent strategies in games with low entropy are more likely to be used in games with high entropy, but not vice versa ([Bednar et al., 2012](#)). In other words, individuals develop strategies for easier games and apply them to more complex games. (2) Another explanation says that expecting others to do what they did in the past (and expecting that they will think you will do what you did in the past, etc.) can coordinate expectations about which of many equilibria will happen ([Devetag, 2005](#)). In other words, players transfer what they did in the past to the subsequent game. (3) Finally, [Knez and Camerer \(2000\)](#) found that transfer of learning across games strongly depended on the presence of superficial, surface similarity (what they call 'descriptive' similarity) between the two games. When the games differed in (what we call) surface characteristics (e.g., actions were numbered differently in the two games) transfer of learning from one game to another did not occur (see a more detailed discussion in [Juvina, Saleem, Martin, Gonzalez, & Lebiere, 2013](#)).

These approaches emphasize factors that pertain to the games (entropy, similarity) or the individuals (expectations). We focus here on factors pertaining to the *interaction* between individuals while not excluding factors related to the game and the individual. We demonstrate that the dynamics of a relational construct – *reciprocal trust* – are key to explaining transfer of learning across games of strategic interaction. Generally, we attempt to bring cognitive-computational and socio-cognitive perspectives into the field of experimental economics, aiming to contribute to theory building and unification.

In the remainder of this paper, we summarize an empirical study on transfer of learning in strategic interaction and present a computational cognitive model as an aid in our attempt to explain the empirical results. We also discuss some of the challenges and opportunities that modeling transfer of learning in strategic interaction brings to the computational cognitive modeling field.

2. Experiment

Only a summary of the experiment is given here; a more detailed description was presented elsewhere ([Juvina et al., 2013](#)). We selected two of the most representative games of strategic interaction: Prisoner's Dilemma (PD) and the Chicken Game (CG). They are both mixed-motive non-zero-sum games that are played repeatedly. The individually optimal and the collectively optimal solutions may be different. Players can choose to maximize short- or long-term payoffs by engaging in defection or cooperation and coordinating their choices with each other. These features give these games the strategic dimension that makes them so relevant to real-world situations ([Camerer, 2003](#)). What makes PD and CG particularly suitable for this experiment is the presence of theoretically interesting similarities and differences, providing an ideal material for studying transfer of learning. [Table 1](#) presents the payoff matrices of PD and CG that were used in this experiment.

Both PD and CG have two symmetric (win-win and lose-lose) and two asymmetric (win-lose and lose-win) outcomes. Besides these similarities there are significant differences between the two

Table 1

Payoff matrices of prisoner's dilemma (PD) and chicken game (CG).

	PD	A	B	CG	A	B
A		−1, −1	10, −10	A	−10, −10	10, −1
B		−10, 10	1, 1	B	−1, 10	1, 1

games. The Nash equilibria are $[-1, -1]$ or $[1, 1]$ ¹ in PD and $[10, -1]$ or $[-1, 10]$ in CG. The number of rounds was not known in advance, so the participants could not apply backward induction. In CG, either of the asymmetric outcomes is more lucrative in terms of joint payoffs than the $[1, 1]$ outcome. This is not the case in PD where an asymmetric outcome $[10, -10]$ is inferior in terms of joint payoffs to the $[1, 1]$ outcome. Mutual cooperation in CG can be reached by a strongly optimal strategy (i.e., alternation of $[-1, 10]$ and $[10, -1]$) or a weakly optimal strategy $[1, 1]$. The optimal strategy in PD corresponds to the weakly optimal strategy in CG numerically, while the strongly optimal strategy of alternation in CG shares no surface-level similarities with the optimal strategy in PD. Thus, although mutual cooperation corresponds to different choices in the two games (i.e., surface-level dissimilarity), they share a deep similarity in the sense that mutual cooperation is, in the long run, superior to competition in both games.

Studying these two games in a sequential ensemble provides a great opportunity to test the theoretical accounts summarized above. Based on the concept of entropy ([Bednar et al., 2012](#)), one would expect transfer of learning to only occur in one direction, that is from PD to CG, because CG has relatively higher entropy (i.e., outcome uncertainty) than PD. According to the "expectation account" ([Devetag, 2005](#)), one would predict that the prevalent strategy from the first game would transfer to the second game. For example, if the two players settle in the $[1, 1]$ outcome in PD, they will be more likely to settle in the $[1, 1]$ outcome in CG as well; if they alternate between the two asymmetric outcomes in CG, they will be more likely to alternate in PD as well. If surface similarities were essential for transfer ([Knez & Camerer, 2000](#)), one would only expect the $[1, 1]$ outcome to drive transfer, because it is identical in the two games.

In contrast, an account focused on interaction would predict that players learn about each other and transfer that learning across games, regardless of surface dissimilarities between games or the order in which games are played.

In both Prisoner's Dilemma and Chicken, learning must occur not only at an individual level but also at a dyad level. If learning occurs only in one of the players in a dyad, the outcomes may be disastrous for that player, because the best solution also bears the highest risk. For example, if only one player understands that alternating between the two moves is the optimal solution in CG, the outcome for that player can be a sequence of -1 and -10 payoffs. Only if both players understand the value of alternation and are willing to alternate, the result will be a sequence of 10 and -1 payoffs for each player, which in average gives each player a payoff of 4.5 points per round. Thus, the context of interdependence makes unilateral individual learning not only useless but also detrimental. The two players must jointly learn that only a solution that maximizes joint payoff is viable in the long term. However, this solution carries the most risk and thus it is potentially unstable in the long term. To ensure that the optimal solution is maintained from one round to another, there must exist a mechanism that mitigates the risk associated with this solution.² It has been suggested that

¹ According to the folk theorem ([Friedman, 1971](#)), the $[1, 1]$ outcome can be a Nash equilibrium if the game is infinitely repeated against the same opponent.

² We do not claim that learning occurs in the two players in the same way or at the same time. It is possible that only one player understands the value of alternating

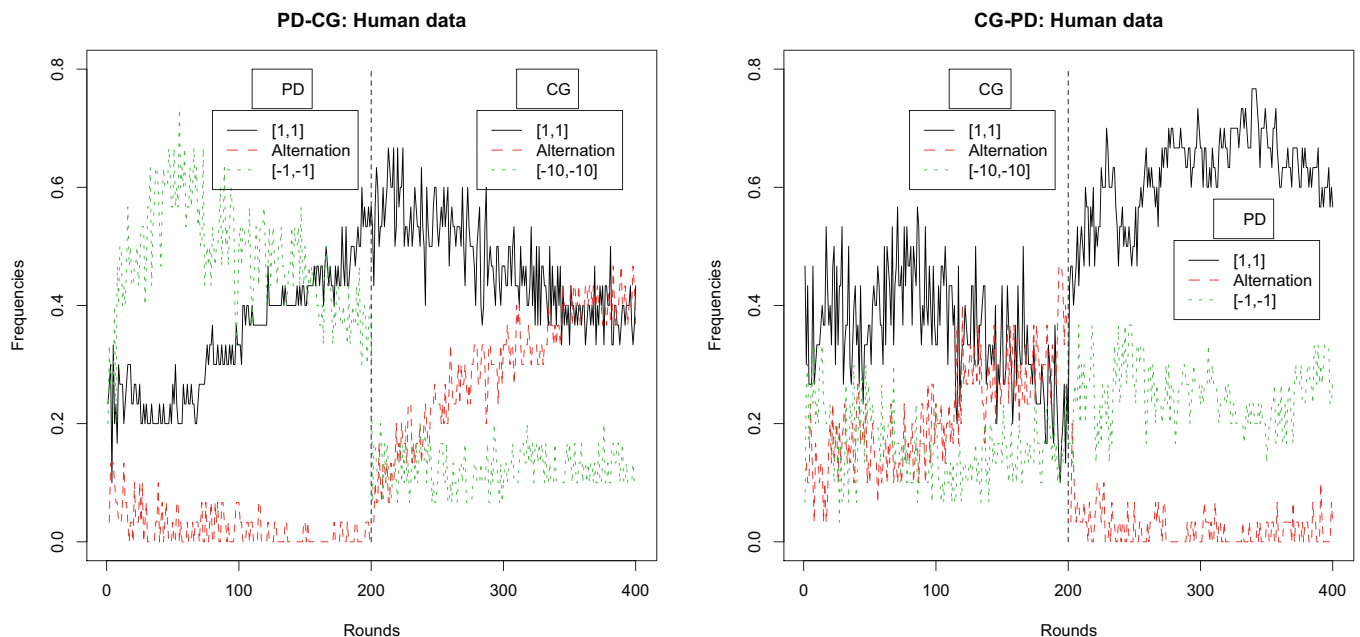


Fig. 1. Frequencies of the most relevant outcomes in PD and CG by order (PD-CG left and CG-PD right) and round averaged across all human participants. Each game was played for 200 rounds.

trust relations are self-sustaining once they have been developed (Hardin, 2002). In situations where there are benefits to individuals that can only be generated through mutual trust, each individual has an incentive to maintain the relation. A trust relation develops through gradual risk-taking and reciprocation (Cook et al., 2005). In turn, as trust develops, risk is reduced and the trust relation becomes more stable. We will demonstrate that the dynamics of reciprocal trust explain specific transfer effects that would not be predicted by the theoretical accounts summarized above.

One hundred and twenty participants were paired with anonymous partners (leading to 60 pairs) and were asked to play the two games in sequence. The 60 pairs were randomly assigned to two conditions defined by the order in which the games were played: PD-CG and CG-PD. Participants played 200 unnumbered rounds of each game. At the end of each game, participants completed a five-item questionnaire assessing: how trustful they were of the opponent; how trustful of them the opponent was; how fair they thought the opponent's actions were; how fair the participants' actions were toward their opponents; and how satisfied they were with the overall outcome of the game.

3. Results³ and discussion

To study transfer of learning across the two games, we analyzed the outcomes of a game according to when it was played. We also analyzed the round-by-round dynamics of these outcomes.

The frequencies of the most relevant outcomes (i.e., the two symmetric ones and an alternation of the two asymmetrical ones

– $[-1,10]$ and $[10,-1]$) are displayed in Fig. 1 on a round-by-round basis. Alternation was defined as 2-round sequences of mutual alternation, that is, the probability that either the sequence “ $[-1,10] \rightarrow [10,-1]$ ” or the sequence “ $[10,-1] \rightarrow [-1,10]$ ” was observed in any two consecutive rounds. We also tested 3- and 4-round sequences: with longer sequences of alternation the data are sparser but the same trends can be observed. The X-axis represents the rounds of the two games: 200 rounds for the first game and 200 rounds for the second game. The Y-axis represents how frequently an outcome was selected. The represented values are averages over 30 pairs of participants. The first thing to notice is how different the two games are from each other from a behavioral perspective: the frequency of actions that lead to the $[1,1]$ outcome (black solid line) increases in Prisoner's Dilemma but decreases in Chicken; alternation (dashed red line) is prominent in CG but almost nonexistent in PD; and the mutually destructive outcome ($[-1,-1]$ in PD and $[-10,-10]$ in CG, dotted green line) is more frequent in PD than in CG. Mutual cooperation is achieved by different strategies in the two games: settling in the $[1,1]$ outcome in PD and alternating between the two asymmetrical outcomes in CG, respectively (see also Bornstein, Budescu, & Zamir, 1997; Rapoport et al., 1976). However, in spite of these differences, mutual cooperation emerges in both games as the preferred solution and it becomes more and more stable over time (see increasing $[1,1]$ curve in PD and alternation curve in CG).

If transfer of learning across games were driven by surface similarities (Knez & Camerer, 2000), one would expect the strategy that is learned in the first game to be applied in the second game as well, even though it may not be appropriate for the second game. This is indeed the case with regard to the $[1,1]$ outcome in the PD-CG order: players learn that $[1,1]$ is long-term optimal in Prisoner's Dilemma and they are more likely to achieve it in the subsequent Chicken Game, even though it is only weakly optimal in Chicken. Fig. 2 shows the frequency of the $[1,1]$ outcome in CG, when CG is played before PD (black solid line) as compared to when CG is played after PD (dashed red line).

In the CG-PD order, if transfer of learning across games were driven by surface similarities, one would expect the strategy of alternating between the two asymmetrical outcomes in CG to be

and decides to act as a strategic teacher (Camerer et al., 2002). Such a player would alternate between the two outcomes with the hope that the other player will eventually cooperate. The second player does not need to perform any high level reasoning for this. As long as she recognizes that the other player is alternating, even if she only myopically best responds to her opponent, she will fall into the optimal alternating behavior. However, there is risk associated with strategic teaching, particularly in Prisoner's Dilemma.

³ Only a summary of the results is provided here as a context for understanding the cognitive model. A more detailed presentation of the empirical results can be found in Juvina et al. (2013).

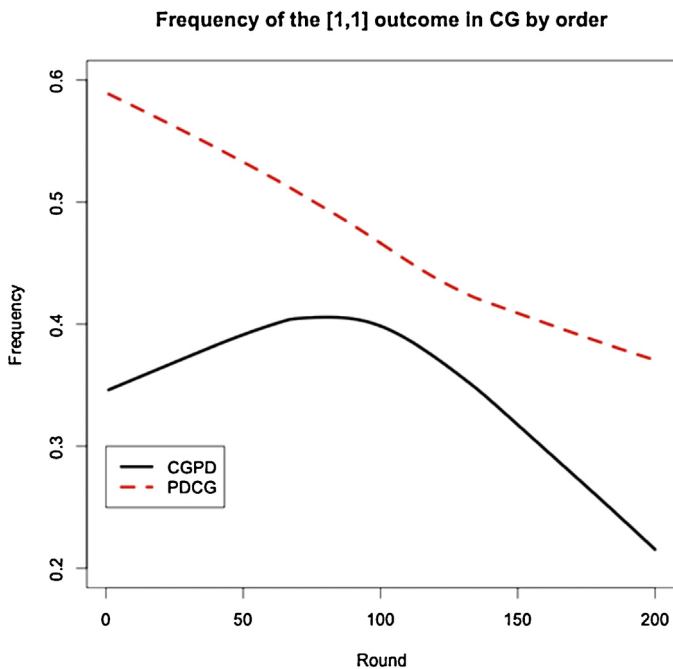


Fig. 2. Transfer of the [1,1] outcome from PD to CG. The frequency of the [1,1] outcome is higher when CG is played after PD (dashed red line) than when CG is played before PD (black solid line). Part of round-by-round variability was removed by smoothing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

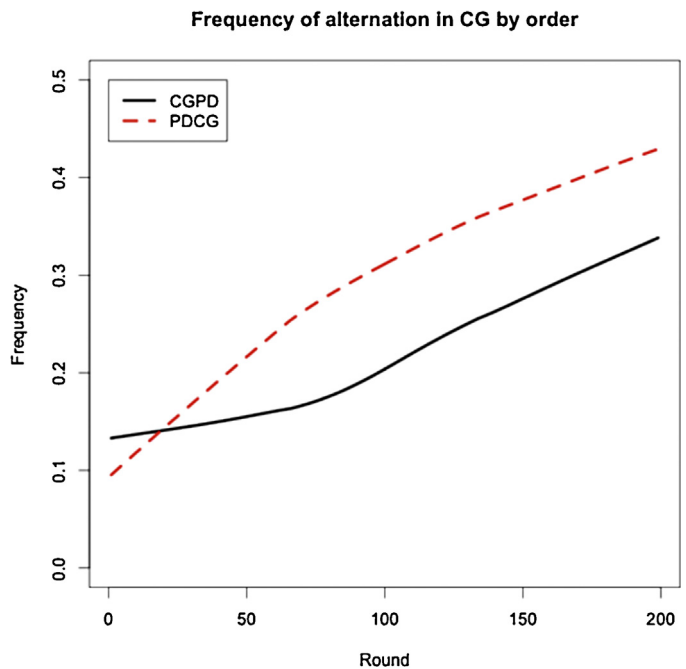


Fig. 3. Transfer of mutual cooperation from PD to CG. The frequency of the alternation outcome is higher when CG is played after PD (red dashed line) than when CG is played before PD (black solid line). Part of round-by-round variability was removed by smoothing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

attempted in PD as well, at least in the beginning of the game. This was not the case (see Fig. 1: alternation in PD is very low regardless of order).

If transfer of learning across games were driven by deep similarities, one would expect learning the optimal strategy in the first game to increase the probability of learning the optimal strategy in the second game, even though there is no surface similarity between these strategies. These strategies ([1,1] in Prisoner's Dilemma and alternation in Chicken) are similar only on an abstract, deep level: they both aim at maximizing joint payoff in a sustainable way, which in these two games is realistically possible only if the two players make (approximately) equal payoffs on a long run. On a surface level, these two strategies are very different. The [1,1] strategy in Prisoner's Dilemma requires that players make the same move at each trial and they do not switch to the opposite move. In contrast, the alternation strategy in Chicken requires that players make opposite moves at each round and they continuously switch between the two moves. Fig. 3 shows a higher level of alternation when CG was played after PD (red dashed line) than when CG was played before PD (black solid line).

The deep transfer effect can be observed in reversed order as well (Fig. 4): Mutual cooperation in PD (the [1,1] outcome) is more frequent when PD is played after CG (red dashed line) than when PD is played before CG (black solid line).

Thus, learning the optimal strategy in the first game increased the probability of learning the optimal strategy in the second game, even though the optimal strategies were different in the two games. This transfer effect was significant in both directions (PD–CG Fig. 3 and CG–PD Fig. 4) contrary to the “entropy” account; if entropy were the causing factor, transfer would have only occurred in one direction – from lower to higher entropy (Bednar et al., 2012).

3.1. Combined effects of surface and deep similarities

In the case of deep transfer, the transfer effect was smaller in magnitude for Chicken than for Prisoner's Dilemma (see

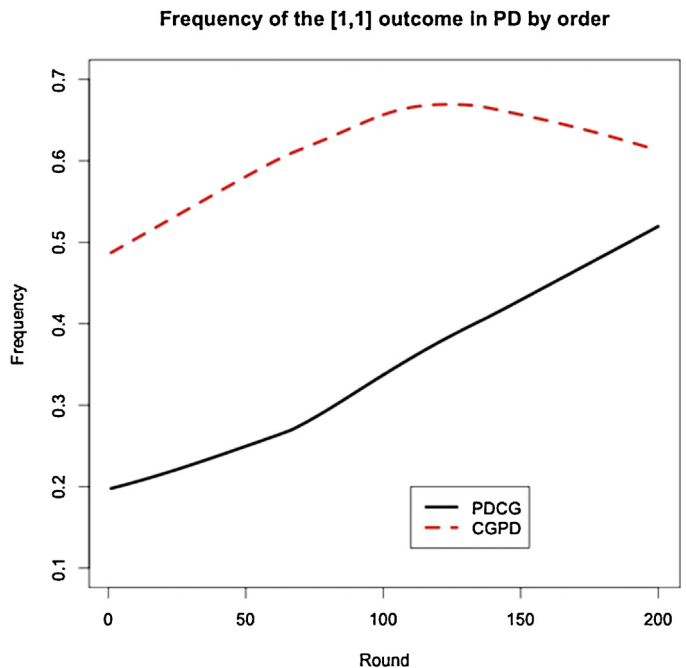


Fig. 4. Transfer of mutual cooperation from CG to PD. The frequency of the [1,1] outcome is higher when PD is played after CG (red dashed line) than when PD is played before CG (black solid line). Part of round-by-round variability was removed by smoothing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Figs. 3 and 4). It seems as if CG has a stronger impact on PD than vice versa. This result further contradicts the entropy account (Bednar et al., 2012), which would predict weak or insignificant transfer from CG (higher entropy) to PD (lower entropy) (CG has higher outcome uncertainty than PD; see Bednar et al., 2012, for a definition of their concept of entropy).

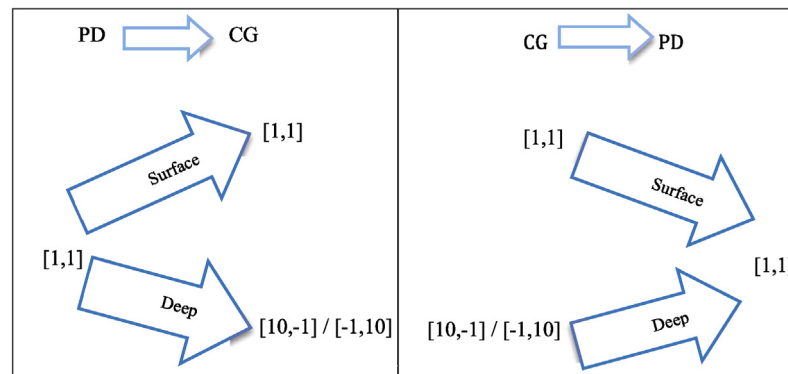


Fig. 5. Combined effects of surface and deep similarities. The left panel shows transfer from PD to CG and the right panel shows transfer from CG to PD. In the left panel, surface and deep transfers are divergent from each other (i.e., they lead to different outcomes). In the right panel, surface and deep transfers are convergent with each other (i.e., they lead to the same outcome). This explains why the observed transfer effect is larger in magnitude in the CG–PD direction than in the PD–CG direction.

The explanation that we propose for this difference is based on how surface and deep similarities combine with each other to drive transfer of learning across games. As a reminder, surface similarities are based on appearance (e.g., [1,1] looks exactly the same in CG and PD) while deep similarities are based on meaning (e.g., [1,1] and alternation maximize joint payoff in PD and CG, respectively). They may have congruent or incongruent effects (Fig. 5). Thus, in the PD–CG order, surface and deep similarities act in a divergent, incongruent way: surface similarity makes it more likely that the [1,1] outcome is selected whereas deep similarity makes it more likely that the alternation outcome is selected. In other words, transfer based on surface similarity interferes with transfer based on deep similarity. In contrast, in the CG–PD order, both types of similarities act in a convergent, congruent way: they both increase the probability that the [1,1] outcome is selected. There is no impeding effect of surface similarity on PD because there is no optimal strategy in CG that is similar enough to a non-optimal or sub-optimal strategy in PD. The impeding and/or enabling effects of surface similarities on deep transfer are revisited in the modeling section.

3.2. Reciprocal trust

In addition to game choices, we analyzed self-reports of reciprocal trust administered at the end of each game. We calculated correlations between these trust variables and the variables indicating mutual cooperation in the two games. We found that the more frequent mutual cooperation was in the first game the more likely the players were to rate each other as trustworthy at the end of the first game. In addition, the more trustworthy players rated each other after the first game, the more likely they were to enact mutual cooperation in the second game. Finally, mutual cooperation in the second game was associated with high levels of trust at the end of the second game. As expected, the level of reciprocal trust increased from the first to the second game. These correlations between trust and the frequency of mutual cooperation suggested that development and maintenance of reciprocal trust facilitated deep transfer of learning across the two games and motivated our modeling approach.

4. A cognitive model of learning and transfer of learning

Modeling transfer of learning across games of strategic interaction provides an opportunity to address some of the ongoing challenges of computational cognitive modeling. Three of these challenges are particularly relevant here and are described below

as the model is introduced. The model is developed in ACT-R and it will be made freely available to the public on the ACT-R website.⁴

ACT-R (Adaptive Control of Thought - Rational) is a theory of human cognition and a cognitive architecture that is used to develop computational models of various cognitive tasks. ACT-R is composed of various modules. There are two memory modules that are of interest here: declarative memory and procedural memory. Declarative memory stores facts (know-what), and procedural memory stores rules about how to do things (know-how). The rules from procedural memory serve the purpose of coordinating the operations of the asynchronous modules. ACT-R is a hybrid cognitive architecture including both symbolic and sub-symbolic components. The symbolic structures are memory elements (chunks) and procedural rules. A set of sub-symbolic equations controls the operation of the symbolic structures. For instance, if several rules are applicable to a situation, a sub-symbolic utility equation estimates the relative cost and benefit associated with each rule and selects for execution the rule with the highest utility. Similarly, whether (or how fast) a fact can be retrieved from declarative memory depends upon sub-symbolic retrieval equations, which take into account the context and the history of usage of that fact. The learning processes in ACT-R control both the acquisition of symbolic structures and the adaptation of their sub-symbolic quantities to the statistics of the environment. ACT-R has been used to develop cognitive models for tasks that vary from simple reaction time experiments to driving a car, learning algebra, and playing strategic games (e.g., [Lebiere, Wallach, & West, 2000](#)).

5. Interdependence

In games of strategic interaction, players are aware of each other and their interdependence. In a previous study we showed that game outcomes were influenced by players' awareness of interdependence. The more information the two players in a dyad had about each other's options and payoffs the more likely they were to establish and maintain mutual cooperation ([Martin, Gonzalez, Juvina, & Lebiere, 2013](#)). Consequently, a cognitive model playing against another cognitive model in a simultaneous choice paradigm needs to develop an adequate representation of the opponent. We use instance-based learning (IBL) ([Gonzalez et al., 2003; Gonzalez et al., in press](#)) to ensure that the opponent is dynamically represented as the game unfolds. Specifically, at each round in the game, an instance (i.e., snapshot of the current situation) is saved

⁴ <http://act-r.psy.cmu.edu/>.

in memory. The instance contains the previous moves of the two players and the opponent's current move. Saved instances are used to develop contextualized expectations about the opponent's moves based on ACT-R's memory storage and retrieval mechanisms (Anderson, 2007). Expectations have been hypothesized to explain some of the spillovers across games (Devetag, 2005).

6. Generality

Before one attempts to build a model of transfer of learning across two games, one needs to have a model that is able to account for the human data in both games. Although by and large cognitive models are task-specific, there is a growing need to develop more general, task-independent models and there are a few precedents: Lebiere, Wallach, and West (2000) developed a model of Prisoner's Dilemma that was able to account for human behavior in a number of other 2×2 games; Lejarraga, Dutt, and Gonzalez (2012) developed an IBL model that made relatively accurate predictions for human behavior in three different binary choice tasks; Gonzalez et al. (in press) have recently expanded this IBL model to account for the dynamics of cooperation in the Prisoner's Dilemma game; and Salvucci (2013) developed a "supermodel" that accounts for human data in seven different tasks. We build upon these precedents of generality by developing a single model to account for round-by-round human data in both PD and CG. We achieve this generality by using instance-based learning for opponent modeling (as described in the previous section) and reinforcement learning for action selection. Both instance-based learning and reinforcement learning are very general learning mechanisms that can produce different results depending on their input. Specifically, at each round in the game, the model anticipates the opponent's move based on the opponent's past behavior and selects its own move based on the utilities of its own past moves in the current context. The input that the model receives as it plays determines the model's behavior. The input is represented by opponent's move, own move, and the payoffs associated with these moves.

An important question is what constitutes the reward from which the model learns the utilities of its actions (moves). Players may try to maximize their own payoff, the opponent's payoff, the sum of the two player's payoffs, the difference, etc. (cf. Gonzalez et al., in press). Thus, a large number of reward structures can be imagined. A complicating assumption is that the reward structure might change as the game unfolds depending on the dynamics of the interaction between the two players. This indeed seems to be the case here, as we have realized after a large number of model explorations: no single preset reward structure is sufficient to account for the human data. One could try to computationally explore the space of all possible reward structures and their combinations to find the one that best fit the human data, but the value of this approach is questionable, because it may lead to a theoretically opaque solution. Instead, we chose to employ a theoretically guided exploration that drastically reduces the number of possible reward structures and, more importantly, gives us a principled way to describe the dynamics of players' motives as the game unfolds (see its description in the next section).

7. Transfer of learning

When the model relies only on the two learning mechanisms described above (i.e., instance-based learning – IBL and reinforcement learning – RL) it is able to only account for the transfer driven by surface similarities. Thus, according to IBL, the opponent is expected to make the same move in a given context as in the previous game. According to RL, an action that has been rewarded in the first game tends to be selected more often in the second game.

It is impossible in this framework to account for transfer driven by deep similarities. For example, if the opponent used to repeat move B when it was reciprocated in PD, there is no reason to switch to alternation between A and B when none of these moves are reciprocated in CG. Moreover, learning within a game may in fact hinder transfer of learning across games if surface similarities are incongruent with the optimal solution in the target game, as in the PD–CG order. To find a solution to the deep transfer problem, we need to return to a theoretical and empirical analysis of the two games.

As mentioned in the introduction, in both Prisoner's Dilemma and Chicken the long-term optimal solution bears the highest risk and, thus, it is unstable in the absence of reciprocal trust. We indeed found that self-reported trust increases after game playing and it positively correlates with the optimal outcome. Recent literature on trust (e.g., Castelfranchi & Falcone, 2010) suggests that trust is essentially a mechanism that mitigates risk and develops through risk-taking and reciprocation. We postulate that trust explains the deep transfer of learning across games. Players learn to trust each other and this affects their reward structure and subsequently their strategies. We added a "trust accumulator" to our model – a variable that increases when the opponent makes a cooperative (risky) move and decreases when the opponent makes a competitive move (see next section for more detail). In addition, another accumulator called "willingness to invest in trust" ("trust-invest accumulator" for brevity) was necessary to overcome situations in which both players strongly distrust each other and persist in a mutually destructive outcome, which further erodes their reciprocal trust, and so on⁵. In such situations, the empirical data shows that players make attempts to develop trust by gradual risk-taking. When these attempts are reciprocated, trust starts to re-develop. In the absence of reciprocation these attempts are discontinued. The trust-invest accumulator increases with each mutually destructive outcome and decreases with each attempt to cooperate that is not reciprocated.

The two accumulators (trust and trust-invest) are used to determine the dynamics of the reward structure. They both start at zero. When they both are zero or negative, the two players act selfishly by trying to maximize the difference between their own payoff and the opponent's payoff. This quickly leads to the mutually destructive outcome, which decreases trust but increases the willingness to invest in trust. When the latter is positive, a player acts selflessly, trying to maximize the opponent's payoff. This can lead to mutual cooperation and development of trust or players may relapse into mutual destruction. When the trust accumulator is positive, a player tries to maximize joint payoff and avoid exploitation. Thus, the model switches between three reward functions depending on the dynamics of trust between the two players. This mechanism provides a principled solution to the problem of selecting the right reward structure and in the same time solves the transfer problem: due to accumulation of trust in the first game, the model employs a reward structure that is conducive to the optimal solution and thus better performance in the second game.

8. Model description

Two ACT-R models run simultaneously and interact with each other. At each round, each model gets as input the game matrix and the opponent model's previous move (as in the human study).

⁵ Procedural learning does not always allow models to escape mutual defection. A mutually destructive outcome can persist in spite of decreasing utilities. For example, in Prisoner's Dilemma, the "defect" rule loses utility from mutual defection, but the "cooperate" rule loses 10 times more from unreciprocated cooperation. This makes the "defect" rule retain relatively higher utility than the "cooperate" rule.

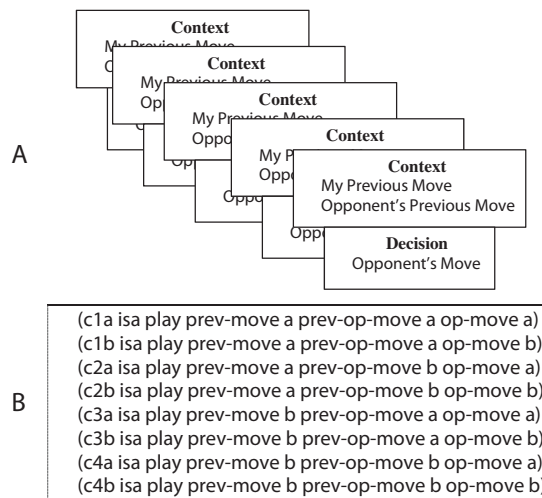


Fig. 6. (A) Instances (snapshots) composed of contexts and decisions are stored in memory. A context is represented by the previous moves of the two players and a decision is represented as the opponent's move in that context. (B) All possible instances that can be used to anticipate the opponent's move: there are four possible contexts and two possible opponent's moves (op-move) in each context.

The first move is random. After the two models make their moves, payoff is assigned based on the payoff matrix.

At each round, the model tries to anticipate the opponent's move based on the opponent's history of moves in similar contexts. In order to learn what move the opponent is likely to make at each round, the model saves instances (snapshots) of prior contexts and the corresponding moves made by the opponent in those contexts (see Fig. 6A). Fig. 6B shows all possible instances that can be encountered during the game. The context is represented by the model's previous move (prev-move) and the opponent's previous move (prev-op-move). In a given context, the opponent can make one of two moves (a or b). For example, the "c1a" instance (the first one in Fig. 6B) is composed of the context "aa" and the decision "a". This instance may be retrieved whenever its context matches the current context of the game. If this instance is retrieved, then the opponent's expected move is "a".

Depending on the opponent's playing history, one of the alternative instances will be more active and more likely to be retrieved from memory. Activation of an instance is a function of the frequency and recency of that instance's occurrence. For example, if in the context "prev-move b prev-op-move a" the opponent usually plays "a", then the instance "c3a" will be more active in the model's memory and thus more likely to be retrieved. Based on retrieval, the model expects that the opponent will play "a" in this context. Anticipation is prone to error due to variability in activations (the ACT-R parameter *activation noise*) and variability in the opponent's behavior. The latter can be caused by the opponent's anticipation uncertainty and strategy shifts (all these sources of variability are independent of each other). Thus, the two models try to anticipate each other's current move based on their respective histories of moves. These anticipations occur in conditions of high uncertainty due to variability of individual model behavior and the context of interdependence. After seeing the actual move of the opponent, the model reinforces the correct instance by rehearsing it, that is, issuing a new retrieval request with the actual opponent's move as a retrieval cue (retrieving an instance increases its activation). This ensures that facts (i.e., actual opponent's moves) are weighted heavier than expectations (i.e., expected opponent's moves), and anticipation errors are not propagated in the long term. Thus, the model leverages the principles of the ACT-R's declarative memory to anticipate the opponent's move.

Table 2

Updating matrices for trust and trust-invest accumulators.

Trust	A	B	Trust-invest	A	B
A	-1,-1	10,-10	A	0.18,0.18	0,-1
B	-10,10	3,3	B	-1,0	0,0

After anticipating the opponent's move, the model must decide on its own move. For example, if the opponent is expected to play "a" the model could decide to play "a" or "b". For this decision, the model leverages the principles of the ACT-R's procedural memory, which is composed of if-then rules. For each possible context (recent moves) and for each possible opponent's move (see Fig. 6), the model contains two decision rules⁶: one that makes the move "a" and one that makes the move "b" (Fig. 7). The first rule "baa-a" can be read as follows: if the goal is to play PD and CG, an instance has been retrieved from memory that matches the current context "ba" and indicates the expected opponent's move to be "a", and the imaginal module is free, then transfer the retrieved instance to the imaginal buffer and make move "a". The second rule "baa-b" can be read as follows: if the goal is to play PD and CG, an instance has been retrieved from memory that matches the current context "ba" and indicates the expected opponent's move to be "a", and the imaginal module is free, then transfer the retrieved instance to the imaginal buffer and make move "b". The retrieved instance is maintained in working memory (imaginal buffer) so that the expected opponent's move (op-move) can be compared with the actual opponent's move.

Each of these rules can fire whenever the context is instantiated and the opponent is expected to make the corresponding move. For example, if the instance "c3a" was retrieved (see the example above), two rules can fire whenever the context is "c3" and the opponent is expected to make move "a": one of these rules makes the move "a" and the other one makes the move "b". Only one rule can fire at a given time, that is, the rule with the higher utility. The two rules start with the same utility, thus they are equally likely to be selected in the beginning of the game. There is random variability in the utilities of the production rules (the ACT-R parameter *utility noise*), which ensures that one of the rules has higher utility than the other. After a move has been made, the model receives the corresponding payoff according to the payoff matrix. The utilities of production rules are updated according to the ACT-R utility learning mechanism (a reinforcement learning algorithm). After a number of rounds, one of the two rules corresponding to a context and an expectation will accrue more utility because it maximizes the reward received by the model. The learning rate of the model (i.e., the ACT-R parameter *alpha*) was fit to match the learning rate observed in the human data.

A key question for this model is what the reward is. If the reward is set to the payoff received from the game matrix, the model cannot account for the deep transfer across games found in the human data (see the section Model validation for a discussion of alternative models). For the reasons discussed above, the rewards are determined by the values of two accumulators: trust and trust-invest. These accumulators are incremented or decremented at each round according to the matrices presented in Table 2. The general trends for the updating of the two trust accumulators are consistent with the theory on trust (e.g., Mayer, Davis, & Schoorman, 1995). Thus, a player's trust increases when the other player cooperates, and more so when the other player unilaterally cooperates, showing willingness to become vulnerable. Conversely, a player's trust decreases when the other player defects and more so when the player's switch to cooperation is not reciprocated. A player's willingness to invest in

⁶ There are 16 decision rules in total.

(p baa-a =goal> isa pd prev-move b prev-op-move a state decide =retrieval> isa play prev-move b prev-op-move a op-move a ?imaginal> state free ==> +imaginal> isa play prev-move b prev-op-move a op-move a =goal> state finish !eval! (setf *move1* 'a) !stop!)	(p baa-b =goal> isa pd prev-move b prev-op-move a state decide =retrieval> isa play prev-move b prev-op-move a op-move a ?imaginal> state free ==> +imaginal> isa play prev-move b prev-op-move a op-move a =goal> state finish !eval! (setf *move1* 'b) !stop!)
--	--

Fig. 7. A couple of production rules can fire in a given context (prev-move = b; prev-op-move = a) and for an expected move of the opponent (op-move = a). The first rule (baa-a) issues move “a” while the second rule (baa-b) issues move “b”. See the text for English renditions of these rules.

trust increases as a function of the need to develop trust, consistent with the idea that trust development activities are more prominent in environments where trust is necessary (e.g., Gambetta & Hamill, 2005; Hardin, 2002). However, a player's willingness to invest in trust development decreases rapidly if a player's switches to cooperation are not reciprocated by the other player, consistent with the idea that trust develops through risk taking and reciprocation (Cook et al., 2005). Aside from these theoretical constraints, the exact values in the updating matrices are free parameters for this model, they were determined by attempting to fit the human data.

At each round, a reward function is selected from a set of three reward functions depending on the sign of the two accumulators as shown in Table 3. When “trust” is positive, regardless of “trust-invest”, the reward function is the sum of the two players' payoffs minus the previous payoff of the opponent. When “trust” is negative and “trust-invest” is positive, the reward function is the opponent's payoff. When both “trust” and “trust-invest” are negative, the reward is the difference between own payoff and the opponent's payoff. The form of the reward functions and their associations with specific values of the two trust accumulators were inspired by the general theory on trust and fine-tuned with the aid of computational exploration. Thus, when trust is positive, the reward function subtracts the previous payoff of the opponent from the current joint payoff. The intuition behind this reward function comes again from the trust theory. Trust mitigates risk in strategic interaction (Cook et al., 2005). In terms of our model, when trust is consistently positive, the associated reward function does not allow a player to consistently make higher payoffs than the other player, which in turn maintains reciprocal trust. In other words, this reward function maintains fairness while maximizing joint payoff (see a more detailed explanation in the General Discussion section). When a model does not trust the opponent (i.e., the trust

accumulator is zero or negative) and is aware of the need to invest in trust development (i.e., the trust-invest accumulator is positive), it tries to maximize the opponent's payoff by unilaterally cooperating, which makes the player vulnerable to exploitation. This decision is justified by the generally accepted definition of trust as a willingness to be vulnerable to the actions of another party (Mayer, Davis, & Schoorman, 1995). If the cooperation attempts are not reciprocated, the trust-invest accumulator is gradually depleted. This makes sense within the trust theory stating that trust relations develop through gradual risk-taking and reciprocation (Cook et al., 2005). When both trust and trust-invest are depleted, the player acts selfishly trying to maximize its own payoff and minimize the opponent's payoff (Table 3).

The trust variables only intervene at key points in the game when reward functions are switched. Between those points the model is not influenced by the trust variables. In such “normal” circumstances, the model behavior is only determined by the current reward function, memory activations, and procedural rules. However, the trust accumulators continue to be updated after each round until they reach the critical values and cause the reward functions to be switched.

One more mechanism was necessary to improve the fit of the model to the human data. Fig. 1 shows that the outcomes in the second game do not start at the chance level (0.25) as in the first game. This suggests that some surface transfer occurs directly without any mediation from the reward function. We modeled this by allowing a third of the model pairs (i.e., 33.33%) to not reset their declarative and procedural memories. This amounts to asserting that these pairs did not notice any fundamental difference between the games and they continued playing as they did in the first game (i.e., they used instances and utilities from the first game to anticipate opponent moves and make moves in the second game).

9. Modeling results

The model described above was fit to the human data presented in Fig. 1. Fitting the model to the human data was done manually by varying a number of parameters (of which some are standard in the ACT-R architecture and others were introduced as part of the trust mechanism) and trying to increase correlation (r) and decrease root mean square deviation (RMSD) between model and human data.

Table 3
The reward function depends on the signs of “trust” and “trust-invest”.

Trust	Trust-invest	Reward function
+	+	Payoff1 + Payoff2 – Prev-Payoff2
+	–	Payoff1 + Payoff2 – Prev-Payoff2
–	+	Payoff2
–	–	Payoff1 – Payoff2

Table 4
Free parameters and their values for the best model fit.

Parameter	Description	Value
Activation noise	Variability in activation of declarative knowledge	0.05
Retrieval threshold	Minimum activation of a retrievable chunk	−5
Latency factor	Determines duration of memory retrievals	0.5
Utility noise	Variability in utility of procedural knowledge	0.02
Learning rate	The rate of learning for procedural knowledge	0.08
Trust increment 1	Trust increment when mutual cooperation	3
Trust increment 2	Trust increment when opponent unilaterally cooperates	10
Trust decrement 1	Trust decrement when mutual defection	−1
Trust decrement 2	Trust decrement when opponent unilaterally defects	−10
Invest increment	Trust invest increment when mutual defection	0.18
Invest decrement	Trust invest decrement when opponent unilaterally defects	−1

Most of the ACT-R parameters were left at their default values and a few were changed from their default but they were maintained fixed across games and rounds. For example, the three parameters mentioned above (activation noise, utility noise, and learning rate) were set at lower values than their defaults. Most important for the model fit were the reward functions and the parameters added as part of the trust mechanism that allowed dynamic changes of the reward function.

The results of the current best model ($r = 0.89$, $\text{RMSD} = 0.09$) are presented in Fig. 8. The model was run 200 times (i.e., 100 pairs).

Overall, the model matches the trends in the human data reasonably well (compare Fig. 8 with Fig. 1). More importantly, the transfer effects are also accounted for. Fig. 9 shows that the [1,1] outcome is more frequent in CG after PD, particularly in the first rounds. The model accounts for this surface transfer by reusing the strategy from the first game. Since this strategy is not the optimal strategy in CG, it is adopted less frequently as the game unfolds.

Fig. 10 shows that alternation is more frequent in CG after PD. The model accounts for this deep transfer by making use of the trust mechanism. The trust accumulated during the first game causes a more cooperative reward function to be used in the second game. Cooperation leads to increased trust, which allows it to continue in spite of being risky.

Fig. 11 shows that the [1,1] outcome is more frequent in PD after CG. This is a case in which deep and surface transfers are congruent with each other (see Fig. 5) and converge to causing a larger transfer effect. The model produces a larger effect (though not quite as large as in the human data) by combining instance based learning, reinforcement learning, and the trust mechanism.

10. Model validation

Since this is a post hoc model (i.e., the model was developed after seeing the human data), model validation can only be incomplete at this moment. We present here three analyses aimed at model validation: split-half cross-validation, model comparison, and individual differences analysis.

We randomly divided the human dataset in two halves and used only the first half to fit the model (training sample). The parameter values for the best model fit are presented in Table 4. Subsequently, the model was compared against the second half of the human data set (testing sample), while maintaining the same parameter values.

Table 5
Model comparison.

Model	Training sample		Testing sample	
	Correlation	RMSD	Correlation	RMSD
Trust	0.87	0.09	0.81	0.11
$P1_n$	0.66	0.19	0.65	0.19
$P1_n - P2_n$	0.01	0.46	−0.001	0.46
$P1_n + P2_n$	0.65	0.29	0.61	0.30
$P1_n + P2_n - P2_{n-1}$	0.66	0.25	0.59	0.27
$P2_n$	0.62	0.34	0.62	0.34

The fit statistics for the training sample are Correlation = 0.87 and Root Mean Square Deviation = 0.09. The prediction statistics for the testing sample are Correlation = 0.81 and Root Mean Square Deviation = 0.11.

To demonstrate that the trust mechanism is necessary to account for the human data we compare the model that includes the trust mechanism (described above) with a number of models that lack such a mechanism. The critical difference between the trust model and the alternative models is the reward function that is used for strategy learning: the trust model dynamically switches between three reward functions depending on trust and trust-invest, whereas each of the alternative models uses only one static reward function. The results of the model comparison are presented in Table 5. The alternative models that are compared against the trust model are labeled according to the reward function from which they learn. $P1_n$ is a model that learns from its own reward. $P1_n - P2_n$ learns to maximize its own payoff and minimize the opponent's payoff. $P1_n + P2_n$ learns to maximize the joint (sum) payoff of the two players, regardless of how the joint payoff is distributed between the two players. $P1_n + P2_n - P2_{n-1}$ subtracts the previous payoff of the opponent from the current joint payoff. This reward function allows the models to learn different cooperation strategies in the two different games (i.e., [1,1] in Prisoner's Dilemma and alternation in Chicken). If the opponent makes the same payoff across consecutive trials, this function defaults to the player's own payoff. However, if the players alternate between low and high payoffs, this function defaults to the opponent's payoff, thus incentivizing the players to continue to alternate. An alternation strategy is learned in games where it is lucrative (e.g., in Chicken, alternating between 10 and −1 yields an average joint payoff of 9) and is not learned in games where it is not lucrative (e.g., in Prisoner's Dilemma, alternating between 10 and −10 yields an average joint payoff of 0). This reward function allowed us to develop one model that reasonably accounts for human behavior in two different games.

As shown in Table 5 the model that includes the trust mechanism outperforms the alternative models for both the training and testing samples. However, the numbers do not tell the whole story. From a qualitative perspective, the difference between the trust model and the alternative models is even clearer. For illustration, Fig. 12 presents the results of a simulation using the best of the alternative models, that is, the model that uses only its own payoff as reward signal ($P1_n$). The outcomes of this model are qualitatively different than the outcomes of the trust model. For example, the most frequent outcome in Prisoner's Dilemma is [1,1] according to the trust model (and the human data) and [−1,−1] according to the alternative model (See Fig. 12 and compare to Figs. 8 and 1).

The model was developed to fit the average human data. However, the human data were spread over a wide range. For example, the level of alternation in Chicken was about 40% by the end of the game. This was the result of averaging the outcomes of pairs that never reached alternation, pairs that alternated only toward the second half of the game, and pairs that alternated throughout the whole game. An analysis of how the model accounts for individual

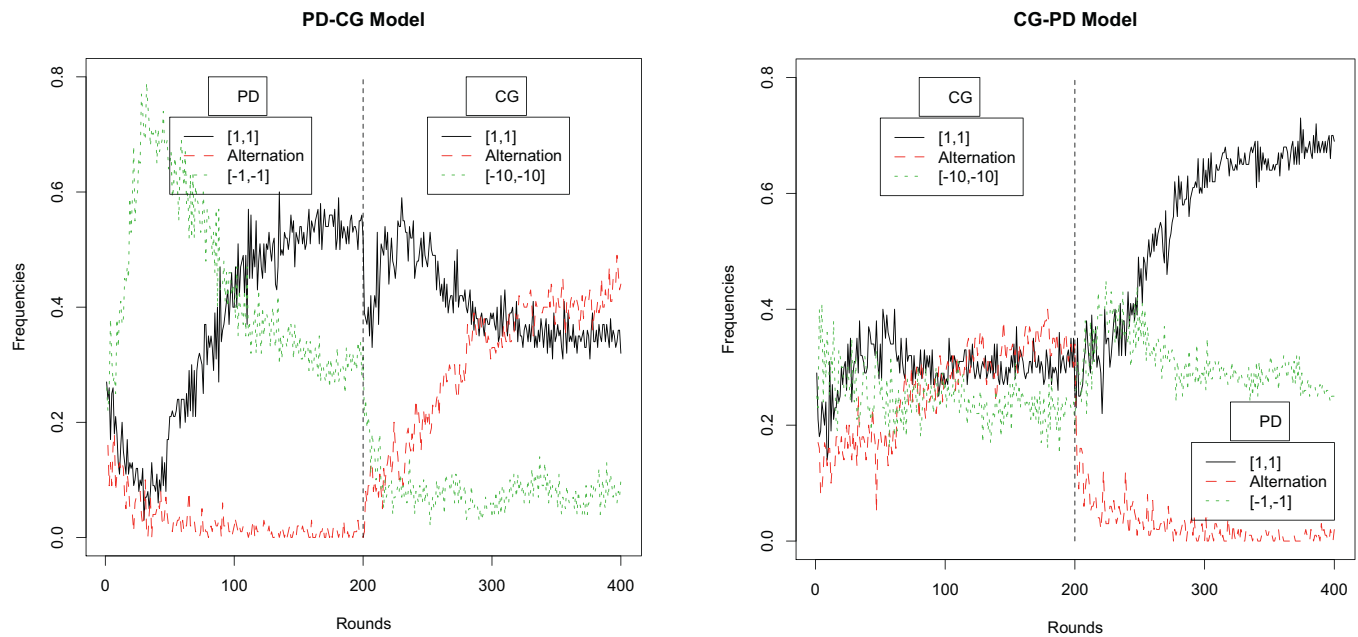


Fig. 8. Results of model simulation. The model reproduces the main trends in the human data (compare to Fig. 1).

differences in the human data can provide additional evidence for model validation. Fig. 13 shows how the human and model results distribute across pairs of individuals. The unit of analysis here is not an individual but a pair of two individuals, because of the interdependence between the two individuals playing a game. The upper row shows Prisoner's Dilemma and the lower row Chicken. The X-axis represents 30 pairs of participants (human and model) and the Y-axis shows the frequency of a particular outcome for each pair. The pairs are ordered according to their frequencies. The human data are represented with solid lines and the model simulations with dashed lines. About 15 pairs achieved the [1,1] outcome in Prisoner's Dilemma (upper left corner) with very low frequency (less than 10%), while other pairs achieved it with higher frequencies (up to 100%). About 5 pairs achieved very high frequencies of the [1,1] outcome (over 90%). The model generates about the same

range of variability, but does not show the polarization seen in the human data, that is, fewer pairs achieve either very low or very high frequencies of the [1,1] outcome. Notwithstanding these limitations, the model reproduces the general pattern of variability in the human data (see all graphs in Fig. 13). This behavior is remarkable, considering that the model was only fit to the average human data. The ability of the model to reproduce this range of individual differences can be attributed to the ACT-R architecture (variability in processing of declarative and procedural knowledge) and the dynamics of the trust mechanism that we introduced here.

These preliminary model validation analyses seem encouraging: the model that we propose here predicts out-of-sample human data reasonably well and capture the qualitative trends in the human data better than alternative models. For model validation to be complete, we will run the model in different conditions and make

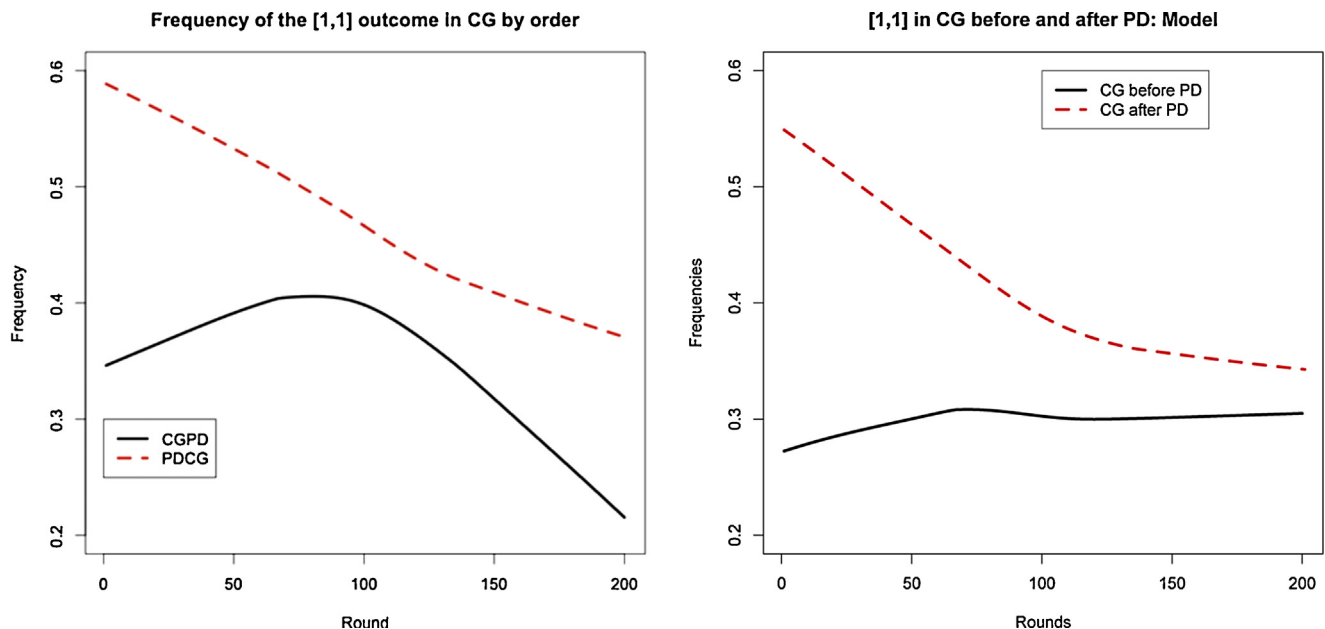


Fig. 9. Comparison of model (right) with human data (left) with regard to the [1,1] outcome in CG.

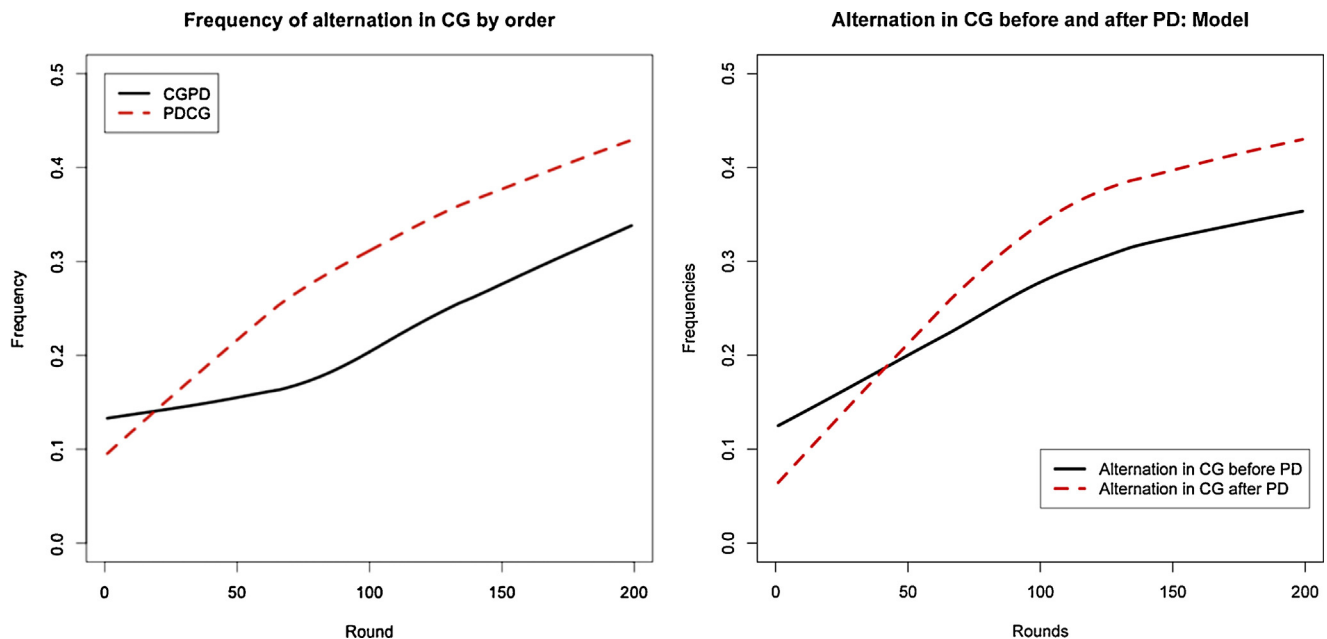


Fig. 10. Comparison of model (right) with human data (left) with regard to alternation.

predictions for a new study. This work will be presented in another paper.

11. General discussion

We employed a cognitive architectural approach to analyze the interplay of different cognitive processes that underlie human behavior in two games of strategic interaction. We developed a single cognitive model to explain human behavior in both games and the transfer of learning from one game to another. We suggest here that a cognitive architectural approach can shed light on phenomena such as gut feelings or intuition, and trust is considered to be one of these phenomena (Gigerenzer, 2007). Much work on intuition and trust remains at an empirical level. Good empirical research is definitely necessary and we contribute to the expanding

area of experimental games. We take a step further and develop a computational cognitive model that has potential theoretical value. We integrate a set of interesting empirical findings in a theoretical and computational framework – ACT-R – that has been developed for decades and applied widely to a variety of paradigms (Anderson, 2007). Cognitive modeling has much to offer to the investigation of strategic decision-making. Pruitt and Kimmel (1977) characterized the field of experimental games as lacking in theory and with little concern for validity. Erev and Roth (1998) argued for the necessity of a Cognitive Game Theory that focuses on players' thought processes and develops simple general models that can be appropriately adapted to specific circumstances, as opposed to building or estimating specific models for each game of interest. In line with this approach, we propose a model that leverages basic cognitive abilities such as anticipating the opponent's move based on

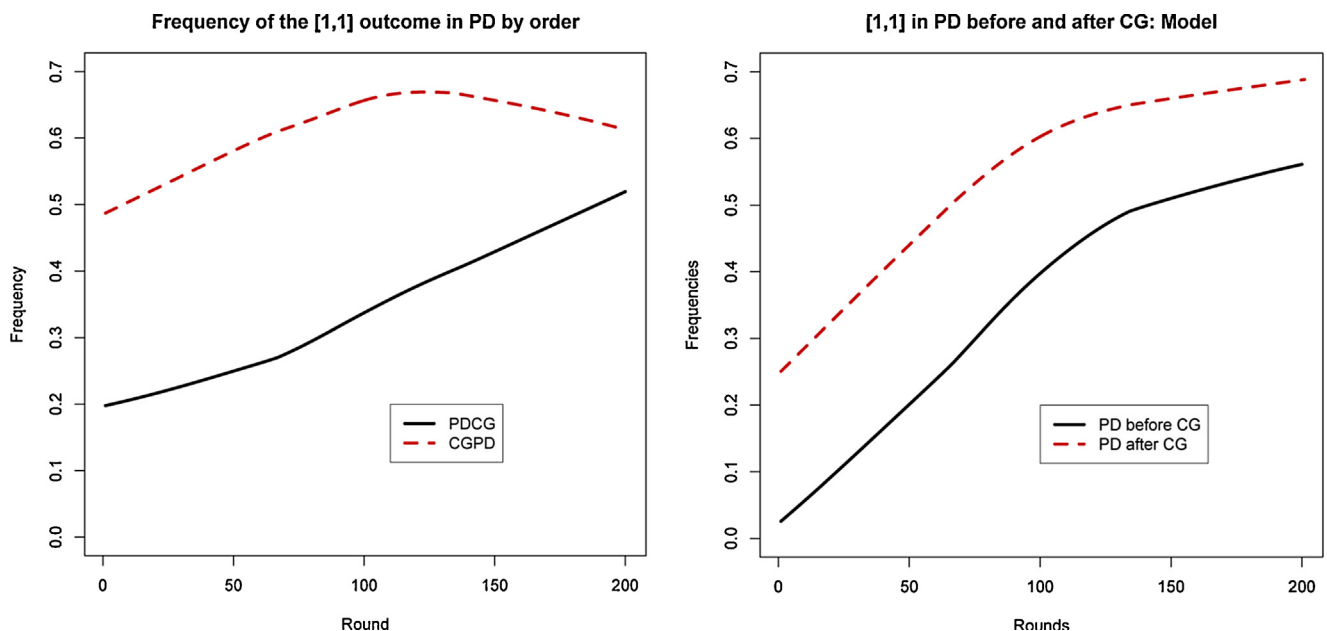


Fig. 11. Comparison of model (right) with human data (left) with regard to the [1,1] outcome in PD.

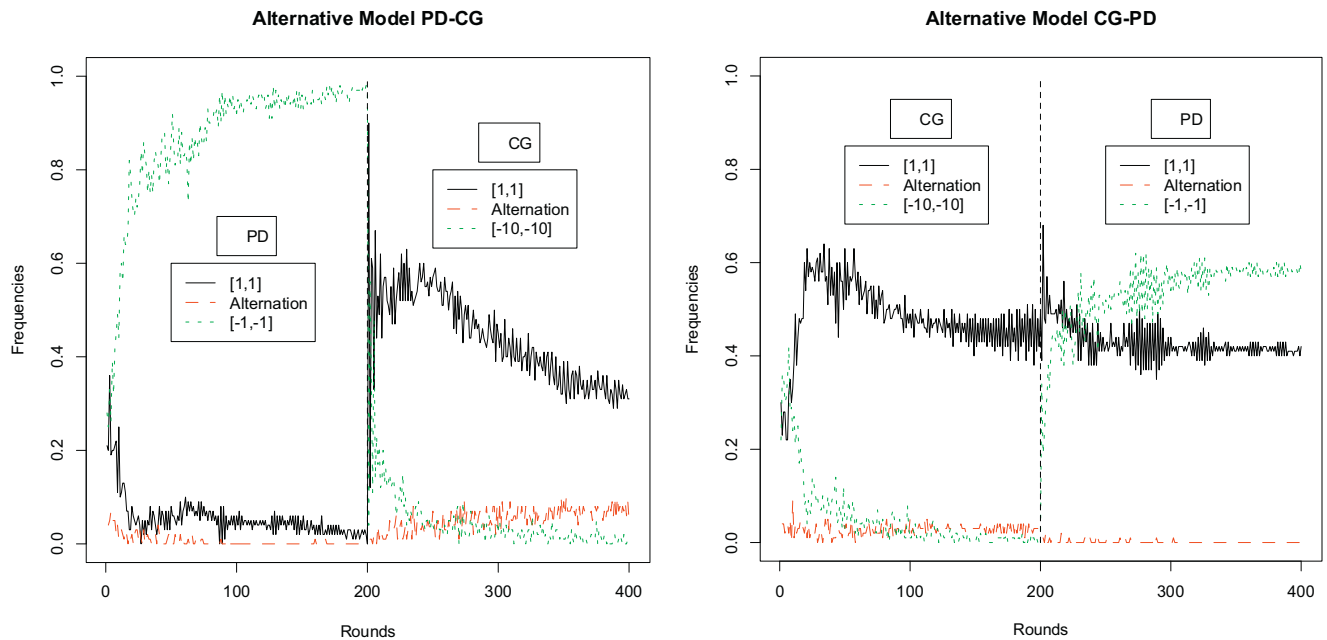


Fig. 12. Results of an alternative model. The alternative model does not reproduce the main trends in the human data as well as the trust model (compare to Figs. 1 and 8).

records of previous moves stored in long-term memory. Memory records include only directly experienced information such as one's own move and the opponent's move. The decision is accomplished by rules that fire depending on their learned utilities. The model

behavior is strongly constrained by learning mechanisms occurring at the sub-symbolic level of the ACT-R cognitive architecture. Our model explains how people make strategic decisions given their experiences and cognitive constraints, without the need for

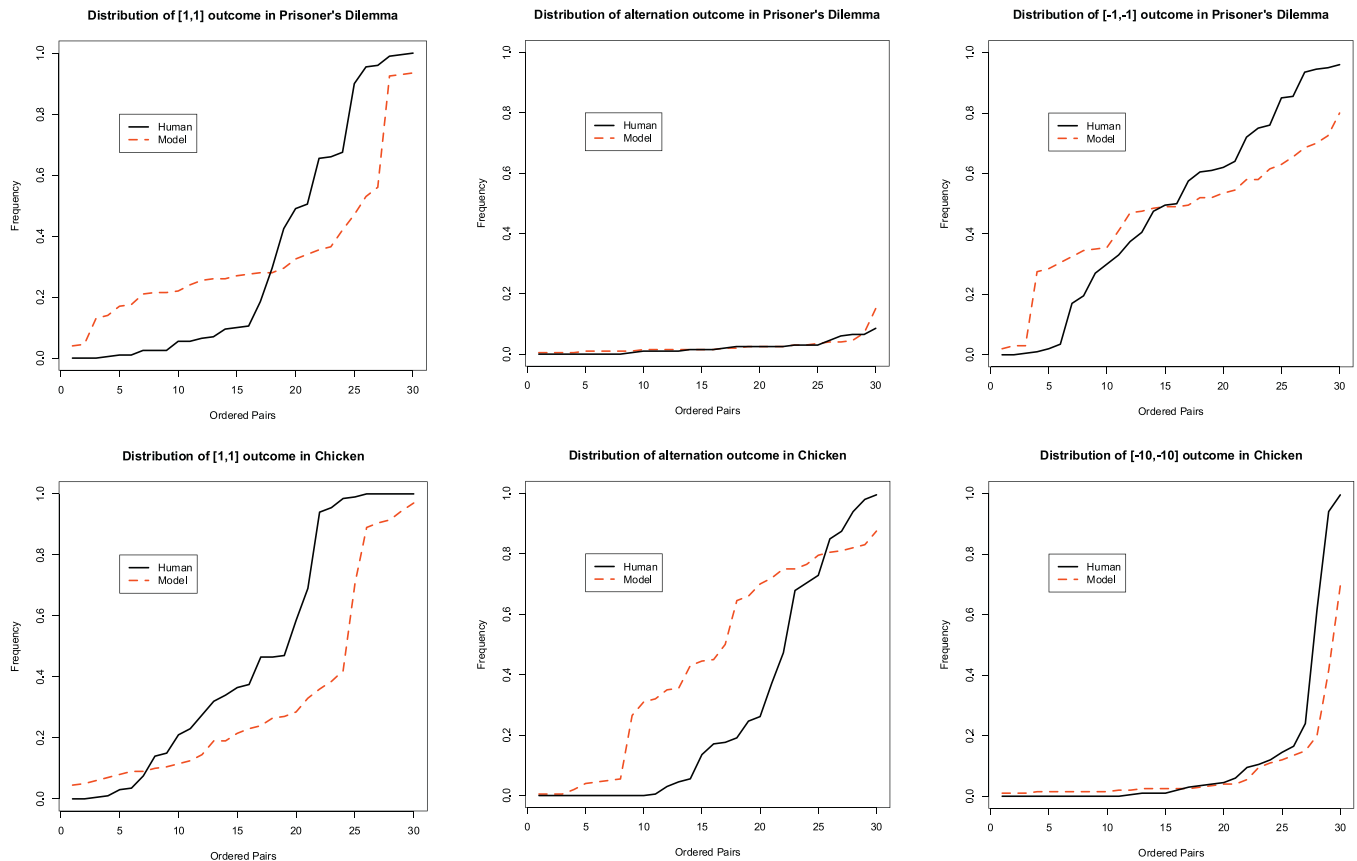


Fig. 13. Distribution of human and model results across pairs of human participants in Prisoner's Dilemma (upper row) and Chicken (lower row). The model reproduces the general pattern of variability in the human data. Some pairs (almost) never achieve mutual cooperation (frequency = 0) while others cooperate for all 200 rounds (frequency = 1).

ad-hoc assumptions. A cognitive architectural approach is sometimes perceived as adding unnecessary complexity. We claim that it adds constraints, generality, and cognitive plausibility. We did not have to reinvent memory, declarative, and procedural learning with all their parameters. Most parameters of the ACT-R architecture were left at their defaults. However, we did have to introduce a trust mechanism to account for the full range of transfer effects with a single model, and this is the main contribution of this work. This mechanism consists of two accumulators, three reward functions, and a set of rules to update the accumulators and switch between reward functions (Tables 2 and 3). This addition was guided by computational exploration and by theory and research on trust and strategic decision-making (Hardin, 2002). Some of our design decisions may seem unusual but they make sense within the established theory on trust. For instance, when a model does not trust the opponent (i.e., the trust accumulator is zero or negative) and is aware of the need to invest in trust development (i.e., the trust-invest accumulator is positive), it tries to maximize the opponent's payoff by unilaterally cooperating, which makes the player vulnerable to exploitation. This decision is justified by the generally accepted definition of trust as a willingness to be vulnerable to the actions of another party (Mayer, Davis, & Schoorman, 1995). If the cooperation attempts are not reciprocated, the trust-invest accumulator is gradually depleted. This makes sense within the trust theory stating that trust relations develop through gradual risk-taking and reciprocation (Cook et al., 2005). Unilateral cooperation with the purpose of achieving mutual cooperation can also be explained by the strategic teaching concept (Camerer, Ho, & Chong, 2002).

Another innovation of our model is the reward function that is used when trust is positive ($P1_n + P2_n - P2_{n-1}$, see Table 3). It subtracts the previous payoff of the opponent from the current joint payoff. This reward function allows the models to learn different cooperation strategies in the two different games (i.e., [1,1] in Prisoner's Dilemma and alternation in Chicken). If the opponent makes the same payoff across consecutive trials, this function defaults to the player's own payoff. However, if the players alternate between low and high payoffs, this function defaults to the opponent's payoff, thus incentivizing the players to continue to alternate. An alternation strategy is learned in games where it is lucrative (e.g., in Chicken, alternating between 10 and -1 yields an average joint payoff of 9) and is not learned in games where it is not lucrative (e.g., in Prisoner's Dilemma, alternating between 10 and -10 yields an average joint payoff of 0). The intuition behind this reward function comes again from the trust theory. Trust mitigates risk in strategic interaction (Cook et al., 2005). In terms of our model, when trust is consistently positive, the associated reward function does not allow a player to consistently make higher payoffs than the other player, which in turn maintains reciprocal trust. In other words, this reward function maintains fairness while maximizing joint payoff. This is a simple way to combine fairness and joint payoff in one criterion, demonstrating how a simple model can learn to exhibit complex behavior.

Could even simpler models explain the behavioral results? The folk theorem postulates that mutual cooperation can be sustained indefinitely in repeated games against the same opponent (Friedman, 1971). Axelrod (1984) showed that cooperation in Iterated Prisoner's Dilemma and other repeated games can be sustained using simple strategies such as tit-for-tat without the need for trust. Simpler models can definitely explain the main trends of the behavioral results, but they say little about the cognitive mechanisms that underlie behavior, such as how people learn the optimal strategies and transfer them to different games. For example, tit-for-tat is insensitive to payoffs and would not be able to switch between [1,1] in Prisoner's Dilemma to alternation in Chicken or vice versa. A more detailed discussion of the necessity of trust was included in the model validation section.

As we wrote this article, we become aware of a number of theoretical and computational approaches that are somewhat similar to our approach. We briefly discuss them here to establish areas of intersection and to highlight the unique contribution of our approach. Rick and Weber (2010) argue that explicit learning of game equilibria is more meaningful and more likely to transfer to similar games than implicit strategy learning that occurs in the presence of feedback. They suggest that limiting or removing feedback altogether promotes explicit learning and deep transfer, whereas providing feedback after each move promotes implicit context-specific learning that fails to transfer. This is perhaps why many studies fail to find deep transfer of learning (as we have also shown in the introduction). We find their argument convincing although in need of refinement and generalization. In our paradigm, the participants learn the games from both description (payoff matrix) and experience (moves they make and payoffs they receive at each round). It would be interesting to disentangle the two sources of learning. Rick and Weber's account would predict that deep transfer would increase in a description-only condition and decrease or disappear in an experience-only condition. We have suggestive evidence that the latter might be true: in a previous study, we found that cooperation decreases in experience-only conditions, where the game matrix is not presented and the participants learn only from feedback (Martin, Gonzalez, Juvina, & Lebiere, 2013). In Rick and Weber's terms, less cooperation implies less meaningful learning available to transfer to other games. Future studies should verify that the descriptive-only condition is indeed more conducive to meaningful learning and deep transfer.

A second approach that is somewhat similar to ours is what Haruvy and Stahl (2012) call "rule learning". They argue that action-based learning cannot account for learning between dissimilar games. "Rule learning entails (1) a specification of behavioral rules, (2) a process for selecting among rules, and (3) a process for updating the likelihood of using these rules" (Haruvy & Stahl, 2012, p. 210). Their rule learning is very similar to some aspects of ACT-R's procedural learning, which entails (1) condition-action pairs (i.e., production rules), (2) conflict resolution, and (3) utility learning. In addition, ACT-R specifies how rules are learned and more generally how procedural learning interacts with declarative learning (belief-learning). In our model, declarative learning is used to develop beliefs about the opponent's moves and procedural learning binds together contexts, beliefs, and actions.

A third approach that resembles ours in some key aspects is the mimicry and relative similarity (MaRS) theory (Fischer et al., 2013). MaRS is a strategy that switches among three states of mimicry – enacted, expected, and excluded – allowing it to select an optimal response for every opponent and stage of the game. Enacted mimicry resembles Tit-For-Tat, expected mimicry resembles Win-Stay-Lose-Shift, and excluded mimicry turns off all cooperative tendencies and transforms MaRS into a defector. The transitions from one state of mimicry to another are guided by the current extent of similarity between the opponent and itself. MaRS monitors two types of similarity by updating two respective registries in a first-in-first-out manner: (i) the passive similarity registry that reflects the proportion of similar choices, regardless of whether they were cooperative or hostile; and (ii) the reactive similarity registry that reflects the opponent's propensity to reciprocate switches toward cooperation that were initiated by MaRS. The former is updated by using values of 1 and 0 for similar and dissimilar choices, respectively. The latter is updated by values of 1 and 0 for reciprocated and non-reciprocated switches, respectively. To assure the genuine nature of reciprocated cooperative switches, MaRS continues to monitor the succeeding moves. If the opponent switches back to defection, the reactive registry is updated with -1, thus canceling out the 1 earned for the previous reciprocation. Continuously monitoring both similarity types allows computing two indices and

applying them as criteria for the selection and transition among the states of mimicry (Fischer et al., 2013). The resemblance of the MaRS strategy to our trust mechanism is remarkable given that they have been developed independently: a first version of our model was presented at a conference in 2012 (Juvina, Lebiere, Gonzalez, & Saleem, 2012). The main point of overlap is the idea of dynamic switches between three strategies depending on the values of two variables that continuously monitor the interaction. Specific to our model is the integration of this mechanism in a cognitive architecture and attempting to fit actual human data. We do not just switch strategies but allow models to learn these strategies in cognitively plausible ways. In addition, we use the exact same model to account for human behavior in two different games and explain transfer of learning between them. We challenge the authors of the MaRS strategy to test it on Chicken without making any changes and more generally to attempt to fit human data in a variety of games.

Lastly, a number of connections can be drawn between our model and the classic social-psychology theories of strategic interaction. Our model's choice is influenced by its expectation about the opponent's choice, although our model's expectations are based on the opponent's past behavior and not on projecting its self-knowledge onto others, as in the social projection theory (Krueger, DiDonato, & Freestone, 2012; Evans & Krueger, 2014). Our model's assumption is that trust is not independent of trustworthiness (Hardin, 2002). Conceivably, in the absence of information about the opponent's past behavior (when trustworthiness cannot be assessed), our model could project its own action tendencies onto the other player. This would be a natural extension to our model that would make it able to fit one-shot social dilemmas. More generally, the interaction between a player's goal and their expectation about the other player's behavior (as in Pruitt and Kimmel's goal/expectation theory, 1977) is analogous to the interaction between declarative and procedural learning in our model. Declarative learning produces contextualized expectations about the other player's behavior and reward functions determine whether the player adopts the goal of maximizing joint payoff or individual payoff. Lastly, our model generates the whole range of individual differences observed in empirical studies; some pairs never achieve mutual cooperation while others cooperate all the time (see Fig. 13). Pair differences arise from variability in the player's cognitive processes and the dynamics of the interaction between players. All models are created equals and they evolve to become cooperators or non-cooperators. They all have the potential to learn to cooperate, but whether their potential is actualized or not depends on the opponent's behavior. In other words, our models are conditional cooperators (Fischbacher, Gächter, & Fehr, 2001). Currently, our model does not assume any a priori differences between players, as in some psychological models that assume, for example, that people are inherently pro-self vs. pro-social (e.g., Bogaert, Boone, & Declerck, 2008). However, our model can easily accommodate such results. For example, individual models can be initialized with different values of the trust variables, reflecting inherent propensities to trust others or be willing to invest in trust development.

Future work should address a number of points. On the empirical side, the trust hypothesis put forth in this paper should be further investigated. For example, in a new experiment, we can randomly rematch the players after the first game. Comparing this treatment to the existing data could provide further empirical evidence in favor of the trust hypothesis. On the modeling side, a validation study is necessary to further test the new set of assumptions about trust that were introduced here. In addition, the model can be further improved by addressing some of its current limitations. For example, the current model does not account for decision time patterns in the human data. It starts with all the necessary declarative and procedural knowledge and only learns the associated sub-symbolic quantities, which is not sufficient to account for

the full range of decision time patterns. However, the ACT-R architecture has a good track record of fitting decision time patterns in a variety of tasks (Anderson, 2007) and we are confident that this can be achieved for the current games as well, by modeling the gradual transition from declarative to procedural knowledge.

12. Conclusion

We developed a computational cognitive model to explain transfer of learning across two games of strategic interaction. One important constraint of our modeling approach was to account for both games with the exact same model. We proposed a simple model that did not have pre-programmed strategies but learned them as it played. The model did not know the specific strategies of the two games; it learned them from experience by making moves and receiving payoffs.

The model explains the observed transfer effects with the aid of a trust mechanism that determines how rewards change depending on the dynamics of the interaction between players. It provides a cognitive and computational account to the human data that challenges the existing theoretical accounts of transfer across games: it demonstrates that transfer occurs in both directions, contrary to what the "entropy account" (Bednar et al., 2012) would predict; it changes its strategy when the game changes, contrary to what the "expectation account" (Devetag, 2005) would predict; and it produces transfer in spite of clear surface dissimilarities, contrary to what the "surface similarity account" (Knez & Camerer, 2000) would predict. We demonstrate that the addition of a relational mechanism – trust – significantly improves the model's match to the human data. We can conclude that factors pertaining to the game or the individual are insufficient to explain the whole range of transfer effects and factors pertaining to the interaction between players should be considered as well.

13. Practical implications

Nowadays, success is increasingly defined based on ability to build cooperating networks and environments, rather than in terms of a win-lose strategy. In particular, in the fields of war and peace-keeping, a strategy of trust development ("winning hearts and minds") is gaining popularity.

The current research on trust is mostly concerned with the general dispositions toward trust and trustworthiness of the actors involved. People are classified as high or low trusters (e.g., Yamagishi, Kikuchi, & Kosugi, 1999) depending on whether they are prone to trust most other people in generic contexts. Trustworthiness is defined as a property of a trusted in relation to an average, generic truster or across a variety of trusters. The trustworthiness of a trusted is assumed to depend more on the ability, benevolence, and integrity of the trusted than on the behavior or the truster (Mayer, Davis, & Schoorman, 1995). However, for many practical purposes it is important to know why and how people become trusting and trustworthy. There is a significant need for studying trust as it develops, erodes, or reemerges in strategic interaction.

Our computational cognitive model of trust dynamics in strategic interaction is a first step toward building a comprehensive theory of trust dynamics grounded in general principles of human cognition. Computational models of this sort can be used to provide insight and specific guidance to decision makers who are interested in trust development within and between organizations. An example of such insight is the need to temporarily focus on maximizing the opponent's payoff in order to restart the trust development process. Based on this insight, service organizations could decide to temporarily focus entirely on their customers' interests in order to gain their trust and loyalty.

Conflict of interest statement

There is no conflict of interest.

Acknowledgments

This research was supported in part by the Defense Threat Reduction Agency (DTRA) grant number HDTRA1-09-1-0053 to Cleotilde Gonzalez and Christian Lebiere and by the U.S. Army Research Laboratory under Cooperative Agreement No. W911NF-09-2-0053. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on. Thanks to the editors and anonymous reviewers for their suggestions.

References

- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.
- Axelrod, R. (1984). *The evolution of cooperation*. Basic Books.
- Bednar, J., Chen, Y., Xiao Liu, T., & Page, S. E. (2012). Behavioral spillovers and cognitive load in multiple games: An experimental study. *Games and Economic Behavior*, 74(1), 12–13.
- Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology*, 47, 453–480.
- Bornstein, G., Budescu, D., & Zamir, S. (1997). Cooperation in intergroup, N-person, and two-person games of chicken. *Journal of Conflict Resolution*, 41(June (3)), 384–406.
- Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, New Jersey: Princeton University Press.
- Camerer, C. F., Ho, T. H., & Chong, J. K. (2002). Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic Theory*.
- Castelfranchi, C., & Falcone, R. (2010). *Trust theory: A socio-cognitive and computational model*. John Wiley and Sons.
- Cook, K. S., Yamagishi, T., Cheshire, C., Cooper, R., Matsuda, M., & Mashima, R. (2005). Trust building via risk taking: A cross-societal experiment. *Social Psychology Quarterly*, 68(2), 121–142.
- Devetag, G. (2005). Precedent transfer in coordination games: An experiment. *Economics Letters*, 89, 227–232.
- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88, 848–881.
- Evans, A. M., & Krueger, J. I. (2014). Outcomes and expectations in dilemmas of trust. *Judgment and Decision Making*, 9(2), 90–103.
- Fischer, I., Frid, A., Goerg, S. J., Levin, S. A., Rubenstein, D. I., & Selten, R. (2013). Fusing enacted and expected mimicry generates a winning strategy that promotes the evolution of cooperation. *Proceedings of the National Academy of Sciences: United States of America*, 110(25), 10229–10233. <http://dx.doi.org/10.1073/pnas.1308221110>
- Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3), 397–404.
- Friedman, J. W. (1971). A non-cooperative equilibrium for Supergames. *Review of Economic Studies*, 38(113), 1–12.
- Gambetta, D., & Hamill, H. (2005). *Streetwise. How taxi drivers establish customers' trustworthiness*. New York: Russell Sage Foundation.
- Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Penguin Books.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gonzalez, C., Ben-Asher, N., Martin, J., & Dutt, V. (2014). A cognitive model of dynamic cooperation with varied interdependency information. *Cognitive Science*, <http://dx.doi.org/10.1111/cogs.12170> (in press)
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27, 591–635.
- Hardin, R. (2002). *Trust and trustworthiness*. New York: Russell Sage Foundation.
- Haruvy, E., & Stahl, D. O. (2012). Between-game rule learning in dissimilar symmetric normal-form games. *Games and Economic Behavior*, 74, 208–221.
- Hertwig, R., & Hoffrage, U. (2013). Simple heuristics: The foundations of adaptive social behavior. In R. Hertwig, U. Hoffrage, & the ABC Research Group (Eds.), *Simple heuristics in a social world*. USA: Oxford University Press.
- Juvina, I., Lebiere, C., Gonzalez, C., & Saleem, M. (2012). Generalization of learning in games of strategic interaction. Paper presented at The Annual Conference of the Cognitive Science Society, Sapporo, Japan.
- Juvina, I., Saleem, M., Martin, J., Gonzalez, M., & Lebiere, C. C. (2013). Reciprocal trust mediates deep transfer of learning between games of strategic interaction. *Organizational Behavior and Human Decision Processes*, 120(2), 206–215.
- Knez, M., & Camerer, C. (2000). Increasing cooperation in prisoner's dilemmas by establishing a precedent of efficiency in coordination games. *Organizational Behavior and Human Decision Processes*, 82, 194–216.
- Krueger, J. I., DiDonato, T. E., & Freestone, D. (2012). Social projection can solve social dilemmas. *Psychological Inquiry*, 23(1), 1–27.
- Lebiere, C., Wallach, D., & West, R. L. (2000). A memory-based account of the prisoner's dilemma and other 2×2 games. Paper presented at the International Conference on Cognitive Modeling.
- Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*, 25(2), 143–153.
- Martin, J. M., Gonzalez, C., Juvina, I., & Lebiere, C. (2013). A description-experience gap in social interactions: information about interdependence and its effects on cooperation. *Journal of Behavioral Decision Making*, <http://dx.doi.org/10.1002/bdm.1810>. Published online in Wiley Online Library (wileyonlinelibrary.com)
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20, 709–734.
- Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28(1), 363–392.
- Rapoport, A., Guyer, M. J., & Gordon, D. G. (1976). *The 2 × 2 game*. Ann Arbor, MI: The University of Michigan Press.
- Rick, S., & Weber, R. A. (2010). Meaningful learning and transfer of learning in games played repeatedly without feedback. *Games and Economic Behavior*, 68(2), 716–730.
- Salvucci, D. D. (2013). Integration and reuse in cognitive skill acquisition. *Cognitive Science*, 37(5), 829–860.
- Thomson, R., Lebiere, C., Anderson, J. R., & Staszewski, J. (2015). A general instance-based learning framework for studying intuitive decision-making in a cognitive architecture. *Journal of Applied Research in Memory and Cognition*, 4(3), 180–190.
- Wegwarth, O., & Gigerenzer, G. (2013). Trust-your-doctor: A simple heuristic in need of a proper social environment. In R. Hertwig, U. Hoffrage, & the ABC Research Group (Eds.), *Simple Heuristics in a Social World*. USA: Oxford University Press.
- Yamagishi, T., Kikuchi, M., & Kosugi, M. (1999). Trust, gullibility, and social intelligence. *Asian Journal of Social Psychology*, 2(1), 145–161.