Evolution of mindsight, transparency and rule-rationality

Rtischev, Dimitry

10 August 2012

Evolution of mindsight, transparency and rule-rationality

Dimitry Rtischev

Faculty of Economics
Gakushuin University
Tokyo, Japan

Working Paper

Revised August 10, 2012

Abstract Evolution of preferences models often assume that all agents display and observe preferences costlessly. Instead, we endogenize mindsight (to observe preferences) and transparency (to show preferences) as slightly costly mechanisms that agents may or may not possess. Unlike in the costless models, we show that universal rule-rationality, mindsight and transparency do not constitute an equilibrium but universal act-rationality, mind-blindness, and opaqueness do. We also find that rule-rationality, mindsight, and transparency may exist in evolved populations, albeit only in a portion of the population whose size fluctuates along an orbit around a focal point. We apply our results to Ultimatum and Trust games to explore how costly and optional mindsight may affect economic performance in interactions among evolved agents.

Keywords evolution of preferences, act-rationality, rule-rationality, ultimatum game, trust game

JEL Classification C73, D83, D87

Acknowledgements I gratefully appreciate comments from seminar attendees at Tsukuba University, Waseda University, and the Institute of Economic Research at Kyoto University, and by participants of the 7th Conference on Logic, Game Theory and Social Choice in Bucharest and the 8th Spain-Italy-Netherlands Meeting on Game Theory in Budapest.

1. Introduction

Most game theory models do not explicitly inquire into the origins of their players. The tacit assumption is that players are creatures of the moment, created out of nothing just to earn the most from playing the game once and then disappear without consequences. By construction, such momentary agents are act-rational in the sense that each player (i) always chooses the action that maximizes his payoff and (ii) assumes that other players also always choose their actions to maximize their payoffs. In the case of the Ultimatum game to divide a resource between a proposer and a responder, a responder who obeys (i) accepts any offer of one cent or more and a proposer who obeys (i) and (ii) offers one cent. But experiments reveal that human subjects usually do not play this way: most proposers make substantial offers and many responders refuse small offers. (Oosterbeek 2004) Evidently many human responders do not obey (i) and most human proposers not obey (i) and/or (ii).

Experimental economics has documented such gaps between theoretical equilibria and experimental play in a variety of games. But considering the different origins of momentary agents and human subjects, the gap in how they play should not surprise. Even anonymous strangers who interact only once in a carefully staged experiment are a product of a long process of evolution. According to evolutionary psychology, humans evolved under selection pressure favoring the ability to make psychological commitments and perceive or infer psychological commitments of others. (Nesse 2001) In particular, evolutionary psychologists stress that human interactions are fundamentally mediated by theory of mind, by subjective commitments secured with emotions not under voluntary control, and by other psychological capabilities refined through selection in social contexts. (Nesse 2001) At the level of the brain, interpersonal neurobiology emphasizes the role of "mindsight," the ability to form and make use of mental representations of how a human (oneself or another) thinks and feels in the midst of an interaction. (Siegel 2001)

A branch of game theory that has gone beyond momentary agents in the direction pointed by evolutionary psychology is evolution of preferences theory, first formulated as the indirect evolutionary approach by Guth and Yaari (1992). Treating agents' decision-making as an endogenous capability subject to evolutionary selection, the theory shows that selection on the basis of relative performance in a strategic interaction yields agents who are committed to pursue a "subjective utility" different from the payoffs they actually earn in the interaction. (Banerjee and Weibull 1995; Guth and Peleg 2001) Asking "What to maximize if you must?" in a generic game, Heifetz, *et. al.* (2007) formally and broadly demonstrated that strategic interaction inherently generates the incentive to commit to maximize something other than the objective payoffs, and that such commitments do not disappear under evolutionary dynamics. In a similar vein, Aumann (2008) informally argues that evolution favors rule-rational agents committed to an optimal rule of behavior over act-rational agents who optimize one act at a time.

All this ultimately rests on Schelling's (1960) insight that if players can make commitments they often find it advantageous to do so, and often the commitments drastically change how a game is played. Since momentary agents in simple games are denied the ability to make commitments, they are act-rational by construction. Since agents in evolution of preferences models are allowed to make commitments, in equilibrium all agents are committed to rule-rational preferences. But this finding of universal rule-rationality among evolved agents critically depends on the assumption that the agents can costlessly display and observe preferences. Attempts to relax this assumption have mostly relied on noisy signaling to model how agents get information about the preferences of other agents. (Guth and Kliemt, 1998 and 2000; Heifetz, *et. al.*, 2007) Instead of stochastically degrading the communication of preferences among agents, we endogenize it by making such communication slightly costly and optional. Specifically, we allow agents to be opaque or transparent, blind or with mindsight, in the following sense: an agent with mindsight observes the preferences of a transparent agent, a blind agent cannot observe another's preferences, and the preferences of an opaque agent are not observable by anyone.

Because mindsight and transparency are slightly costly in our model, the rule-rational population in which all agents are transparent and have mindsight can be invaded by blind and opaque agents. Reversing the findings of the costless evolution of preferences models, we show that universal rule-rationality is not an equilibrium but universal act-rationality is. However, we also find that there exists a focal point surrounded by closed orbits along which rule-rational transparent agents and agents with mindsight may exist in in the long-run and in significant proportions. We apply our results to Ultimatum and Trust games to explore how economic performance in evolved populations is affected by rule-rationality, mindsight, and transparency.

The distinction between rule-rationality and act-rationality has been long noted by philosophers of rationality and morality, albeit using different terminology. (Gauthier, 1986, Chapter VI) Of particular relevance to our model is Danielson's (1990) pioneering book which explicitly considers mindsight and transparency among strategically interacting agents. Attempting to algorithmically examine Gauthier's (1986) theory of rational morality, Danielson conceives agents as logic programs that may examine other agents' programs and may allow themselves to be examined by others' programs. Like Danielson, we assume each agent operates according to its own built-in "decision logic" and allow some agents to display and observe these decision logics. Unlike Danielson, we assume that the display and observation of decision logics are costly and focus on evolutionary population dynamics.

The rest of the paper is organized as follows. The next section lays out the formal framework. Section 3 presents results of equilibrium analysis. We then explore our findings in the context of the Ultimatum Game (Section 4) and the Trust Game (Section 5). Section 6 concludes.

2. The model

There are two separate populations of agents: "leaders" and "responders." A dyad is formed by randomly drawing one agent from each population. The state of the environment (*e.g.,* weather) that prevails when the dyad interacts is expressed by random variable *e*, which is drawn every time a dyad is formed.[1] Each dyad plays a base game as follows: first both players observe *e*, then the leader observes the responder and takes action *x*, and finally the responder observes *x* and takes action *y*. The resulting payoffs are $\pi_1(x, y, e)$ to the leader and $\pi_2(x, y, e)$ to the responder. When a dyad is formed, the leader is endowed with π_1^0 and the responder with π_2^0 . If at least one player abstains from the game, both players keep these endowments.

We make the following assumptions about the cognitive structures and processes that agents use to choose actions.

Definition *Decision logic* is a deterministic function, algorithm, or program within an agent that computes the action the agent takes whenever he plays.

Assumption Each responder has a decision logic whose inputs are the leader's action and the state of the environment: $y = \Theta_i(x, e)$.

Definition *Theory of mind* is what a leader believes to be the decision logic of a responder in his dyad.

Assumption Each leader has a decision logic whose inputs are a theory of mind and the state of the environment: $x = \Lambda(\Phi_i, e)$.

Assumption Each leader's decision logic maximizes the payoff given the theory of mind:

$$\Lambda(\Phi, e) = \arg\max_x \pi_1(x, \Phi(x, e), e) .$$

Every leader is one of two psychological types: either blind (type B) or with mindsight (type M). Every responder is one of two psychological types: either transparent (type T) or opaque (type O). We define these properties as follows.

Definition The decision logic of a leader with *mindsight* takes as input the decision logic of a transparent responder.

[1] Although *e* does not play a direct role in the analysis, it is necessary to assume a changing environment in order to ensure that agents dynamically compute their actions rather than being hardwired.

Definition The decision logic of a *transparent* responder is taken as input by the decision logic of every leader with mindsight.

Definition The decision logic of a *blind* leader cannot take as input the decision logic of any responder.

Definition The decision logic of an *opaque* responder cannot be taken as input by any leader.

In terms of psychological traits, there are four possible types of dyads: MT, MO, BT, and BO. In the MO, BT, and BO dyads, the leader cannot observe responder's decision logic and therefore the leader's decision logic relies on a built-in theory of mind Φ which may or may not be accurate. Only in the MT dyad the leader observes the responder's decision logic and uses it as the theory of mind.[2]

The type of responder is given by (θ, Θ), where $\theta \in \{T, O\}$ indicates transparency or opaqueness and Θ is the decision logic; that is $y = \Theta(x, e)$. Displaying one's decision logic is a costly capability – a transparent responder incurs a cost $\tau > 0$ every time he plays. Many types of opaque and transparent responders may exist, differing in terms of their decision logic. The state of responder population is given by the population share vector[3] $\mathbf{q} = (q_1, ..., q_O, q_{O+1}, ..., q_{O+T})$, where $q_i \in [0,1]$ is the share of the i^{th} type of responder (θ_i, Θ_i), O is the number of opaque responder types, T is the number of transparent responder types, and $\sum q_i = 1$.

The type of leader is specified by (λ, Φ), where $\lambda \in \{B, M\}$ indicates blindness or mindsight and Φ is the theory of mind. Mindsight is a costly capability – a leader with mindsight incurs a cost $\mu > 0$ every time he plays. Many types of blind leaders and leaders with mindsight may exist, differing in terms of their decision logic. The state of leader population is given by the population share vector $\mathbf{p} = (p_1, ..., p_B, p_{B+1}, ..., p_{B+M})$, where $p_i \in [0,1]$ is the share of the i^{th} type of leader (λ_i, Φ_i), B is the number of blind leader types, M is the number of leader types with mindsight, and $\sum p_i = 1$.

The payoffs earned in each of the four dyads are as follows[4]:

Blind-Opaque dyad: (leader type $i = 1, ..., B$; responder type $j = 1, ..., O$)

$$\Pi_{ij}^{BO} = \pi_1(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) \text{ to the leader}$$

---

[2] In general, the leader in the MT dyad could ignore the input and use an incorrect theory of mind. As proven in Proposition 2, doing so cannot increase the leader's payoff in the dyad.

[3] Population share vectors **q** and **p** are column vectors.

[4] For notational clarity, we omit the environment parameter *e* in most expressions hereinafter.

$$\Pi_{ij}^{OB} = \pi_2(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) \text{ to the responder}$$

Blind-Transparent dyad: ($i = 1,\ldots,B$; $j = O+1,\ldots,O+T$)

$$\Pi_{ij}^{BT} = \pi_1(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) \text{ to the leader}$$

$$\Pi_{ij}^{TB} = \pi_2(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) - \tau \text{ to the responder}$$

Mindsight-Opaque dyad: ($i = B+1,\ldots,B+M$; $j = 1,\ldots,O$)

$$\Pi_{ij}^{MO} = \pi_1(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) - \mu \text{ to the leader}$$

$$\Pi_{ij}^{OM} = \pi_2(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) \text{ to the responder}$$

Mindsight-Transparent dyad: ($i = B+1,\ldots,B+M$; $j = O+1,\ldots,O+T$)

$$\Pi_{ij}^{MT} = \pi_1(\Lambda(\Theta_j), \Theta_j(\Lambda(\Theta_j))) - \mu \text{ to the leader}$$

$$\Pi_{ij}^{TM} = \pi_2(\Lambda(\Theta_j), \Theta_j(\Lambda(\Theta_j))) - \tau \text{ to the responder}$$

The leaders' payoff matrix has $B+M$ rows and $O+T$ columns arranged as follows:

$$\mathbf{\Pi}_L = \begin{bmatrix} \mathbf{\Pi}^{BO} & \mathbf{\Pi}^{BT} \\ \mathbf{\Pi}^{MO} & \mathbf{\Pi}^{MT} \end{bmatrix}$$

where $\mathbf{\Pi}^{BO} = [\Pi_{ij}^{BO}]$ is the B-row O-column matrix of leader payoffs in blind-opaque dyads, $\mathbf{\Pi}^{BT}$ is the B-row T-column matrix of leader payoffs in blind-transparent dyads, $\mathbf{\Pi}^{MO}$ is the M-row O-column matrix of leader payoffs in mindsight-opaque dyads, and $\mathbf{\Pi}^{MT}$ is the M-row T-column matrix of leader payoffs in mindsight-transparent dyads.

Analogously, the responders' payoff matrix has $B+M$ rows and $O+T$ columns arranged as follows:

$$\mathbf{\Pi}_R = \begin{bmatrix} \mathbf{\Pi}^{OB} & \mathbf{\Pi}^{TB} \\ \mathbf{\Pi}^{OM} & \mathbf{\Pi}^{TM} \end{bmatrix}$$

The evolutionary dynamics occur as follows. During each generation many random dyads are formed to play the base game. Each type of leader (responder) accumulates fitness equal to the sum of the payoffs earned by that type of leader (responder) in the base game. At the end of a generation agents replicate and die. Replication occurs within the leader and responder populations separately. The replication is governed by a standard replicator dynamic. Specifically, the share of a given type of leader (responder) in the new population of leaders (responders) equals the fitness share earned by that type of leader (responder) in the old population, computed as the share of the total fitness earned by all leaders (responders) in the old generation. Many generations ensue.

The expected fitness of each leader type given the state of the responder population is given by the expected fitness vector $\mathbf{V}^L \equiv \mathbf{\Pi}_L \mathbf{q}$. The population average fitness of leaders is $\overline{V}_L \equiv \mathbf{p} \cdot \mathbf{V}^L$. Analogously, the expected fitness of each responder type given the state of the leader population is $\mathbf{V}^R \equiv \hat{\mathbf{\Pi}}_R \mathbf{p}$, where $\hat{\mathbf{\Pi}}_R$ is the transpose of $\mathbf{\Pi}_R$. The population average fitness of responders is $\overline{V}_R \equiv \mathbf{q} \cdot \mathbf{V}^R$. The replicator dynamic is:

Leaders: $\dot{p}_i = p_i(V_i^L - \overline{V}_L), \quad i = 1,...,B+M$

Responders: $\dot{q}_j = q_j(V_j^R - \overline{V}_R), \quad j = 1,...,O+T$

Definition A *fixed point* is a population state of leaders and responders (**p**,**q**) that satisfies the following conditions for all *i = 1,…, B+M* and *j = 1,…, O+T*:

(i) $V_i^L = \overline{V}_L \quad if \ p_i > 0$

(ii) $V_j^R = \overline{V}_R \quad if \ q_j > 0,$

(iii) $V_i^L < \overline{V}_L \quad if \ p_i = 0$

(iv) $V_j^R < \overline{V}_R \quad if \ q_j = 0$

In our framework there are two populations of agents, each playing a different role in an asymmetric base game. As is commonly done in evolutionary game theory models, we could have introduced a stage in which role assignment is randomly determined before play in each dyad. This would have made the base game symmetric and allowed all the players to belong to a single population. The benefit of our approach is that it retains separate focus on mindsight and transparency, which are inherently asymmetric capabilities for receiving and transmitting information about preferences.

3. Equilibrium analysis

Although the space of possible decision logics is very large, two decision logics play a central role in equilibrium analysis. Following Aumann (2008), we will refer to them as act-rationality and rule-rationality.

Definition An *act-rational* responder has decision logic $A(x,e) = \arg\max_y \pi_2(x,y,e)$.

Definition A *rule-rational* responder has decision logic *R(x, e)* such that

$$\forall e \quad \nexists R' \; s.t. \; \pi_2(\Lambda(R'), R'(\Lambda(R')), e) > \pi_2(\Lambda(R), R(\Lambda(R)), e) \quad \text{and}$$

$$\forall e \quad \forall x \neq \Lambda(R) \quad \nexists R' \; s.t. \; \pi_1(x, R'(x), e) > \pi_1(x, R(x), e)$$

Decision logic A is the responder's best response whereas decision logic R is the responder's best strategic commitment. The first condition in the definition of R ensures that no other decision logic yields the responder a higher payoff. The second condition ensures R "punishes" as much as possible a leader who does not maximize own payoff by taking into account the leader's commitment to R.

As a shorthand, we will denote the leader's payoff-maximizing strategy given the decision logic of the responder as follows:

$$x_A \equiv \Lambda(A) = \arg\max_x \pi_1(x, A(x))$$

$$x_R \equiv \Lambda(R) = \arg\max_x \pi_1(x, R(x))$$

We will also use the following shorthand notation to denote base game payoffs to leaders (*i=1*) and responders (*i=2*), gross of the costs of mindsight and transparency:

$$\pi_i^{RR} = \pi_i(x_R, R(x_R))$$
$$\pi_i^{AA} = \pi_i(x_A, A(x_A))$$
$$\pi_i^{RA} = \pi_i(x_R, A(x_R))$$
$$\pi_i^{AR} = \pi_i(x_A, R(x_A))$$

We confine attention to base games in which strategic commitment affects payoffs. This class of games is large and can be formally described as in Heifetz, *et. al.* (2007). For our purposes, it suffices to assume the following about the payoff structure of the base game:

Assumption 1 The base game is such that x_A, and x_R are uniquely defined and satisfy the following:

$x_A \neq x_R \; \forall e$

$\pi_i^{AA} > \pi_i^0 \; \forall e$ (participation constraint under act-rationality)

$\pi_i^{RR} > \pi_i^0 \; \forall e$ (participation constraint under rule-rationality)

We also assume that the costs of mindsight and transparency are not too large relative to their effect on payoffs in the base game. Specifically:

Assumption 2 For a responder facing a leader with mindsight, the cost of transparency is less than the benefit of strategic commitment:

$$\tau < \pi_2^{RR} - \pi_2^{AA} \quad \forall e$$

Assumption 3 For a leader facing a rule-rational transparent responder, the cost of mindsight is less than the benefit of heeding the responder's strategic commitment:

$$\mu < \pi_1^{RR} - \pi_1^{AR} \quad \forall e$$

The following propositions characterize fixed point populations. All the proofs are in the Appendix.

Proposition 1 At a fixed point, every opaque responder is act-rational.

Proposition 2 At a fixed point, every leader with mindsight uses the transparent responder's decision logic as the theory of mind and believes that an opaque responder is act-rational. That is, a leader with mindsight in a dyad with a responder of type (θ, Θ) has the theory of mind:

$$\Phi^M = \begin{cases} \Theta \; if \;\; \theta = T \\ A \; if \;\; \theta = O \end{cases}$$

Proposition 3 The only monomorphic population that is a fixed point is: all responders are opaque and act-rational, all leaders are blind and hold act-rationality as the theory of mind. Moreover, this is a stable fixed point.

Proposition 4 A population in which all leaders have mindsight is not a fixed point.

Proposition 5 A population in which all responders are transparent and have the same decision logic is not a fixed point.

Proposition 6 If the cost of mindsight is sufficiently small, there exists a unique fixed point at which a share $m^* \in (0,1)$ of leaders have mindsight, a share $t^* \in (0,1)$ of responders are transparent, all blind leaders believe responders are rule-rational, and all transparent responders are rule-rational. This fixed point is given by

$$m^* = \frac{\pi_2^{RR} - \pi_2^{RA} - \tau}{\pi_2^{AA} - \pi_2^{RA}} \qquad t^* = 1 - \frac{\mu}{\pi_1^{AA} - \pi_1^{RA}}$$

Moreover, *(m*, t*)* is an evolutionary focal point around which all trajectories are closed orbits with the time frequencies of *(m, t)* along the orbits equal to *(m*, t*)*.

Proposition 7 There exists a unique unstable fixed point at which a fraction $m^* \in (0,1)$ of leaders have mindsight, a fraction $t^* \in (0,1)$ of responders are transparent, all blind leaders believe responders are act-rational, and all transparent responders are rule-rational.

The foregoing propositions identify combinations of psychological traits and decision logics that may evolve among agents subject to selection based on their performance in a sequential dyadic interaction. The only monomorphic evolutionarily stable state is universal blindness, opaqueness, and act-rationality. Blind leaders and opaque responders cannot go extinct. Opaque responders are act-rational. Although neither a blind leader nor a leader with mindsight can see the decision logic of an opaque responder, the leader with mindsight has an advantage when paired with an opaque responder because by seeing opaqueness he can infer that the responder is act-rational. Although mindsight, transparency, and rule-rationality cannot be universal and cannot be present in asymptotically stable proportions, they may nevertheless be present in an evolved population along closed orbits around a focal point. In such populations, blind leaders assume that responders are rule-rational and leaders with mindsight assume that opaque responders are act-rational.

4. Ultimatum island

Imagine an island populated by two species: pushers and pullers. Every minute, a random pusher finds a resource of value *e* and a random puller appears nearby. To extract the resource they must cooperate: the pusher must push while the puller pulls. The pusher offers to give *x* to the puller after they extract the resource, where $0 < \varepsilon \le x \le e - \varepsilon$ and ε is the minimum amount that can be allocated to an agent.[5] The puller accepts or rejects the offer. If the puller rejects, the agents go their separate ways and the resource rots away. If the puller accepts, the agents cooperate to extract and divide the resource. There are no endowments that agents can keep by abstaining from the game: $\pi_1^0 = \pi_2^0 = 0$. The pusher and puller payoff functions are, respectively

$$\pi_1(x,y) = \begin{cases} e - x & if\ \ y = accept \\ 0 & if\ \ y = reject \end{cases} \qquad \pi_2(x,y) = \begin{cases} x & if\ \ y = accept \\ 0 & if\ \ y = reject \end{cases}$$

Under ideal institutions that costlessly ensure cooperation in all dyads, realized product is $P = e$. This is the first-best baseline. We ask: What kind of rationality may evolve on this island?

The decision logic of an act-rational puller is:

$$A(x) = \begin{cases} accept\ if\ \ x \ge \varepsilon \\ reject\ if\ \ x < \varepsilon \end{cases}$$

The decision logic of a rule-rational puller is:

$$R(x) = \begin{cases} accept\ if\ \ x \ge e - \varepsilon \\ reject\ if\ \ x < e - \varepsilon \end{cases}$$

Pusher strategies are: $x_A = \varepsilon$ and $x_R = e - \varepsilon$. The payoffs under the various combinations of decision logics are:

$$\begin{array}{ll} \pi_1^{AA} = e - \varepsilon & \pi_2^{AA} = \varepsilon \\ \pi_1^{RR} = \varepsilon & \pi_2^{RR} = e - \varepsilon \\ \pi_1^{AR} = 0 & \pi_2^{AR} = 0 \\ \pi_1^{RA} = \varepsilon & \pi_2^{RA} = e - \varepsilon \end{array}$$

By Proposition 3, the blind/opaque/act-rational population in which all pushers are (B, A) and all pullers are (O, A) is evolutionarily stable.

[5] To avoid weak inequalities, we assume that when cooperating each agent incurs a small cost and therefore agents make and accept only those offers which give both parties a strictly positive gain of at least ε.

By Proposition 6, the following population is an evolutionary focal point:

Pushers: (B, R) and (M, Φ^M) Population shares: $b = 1 - \frac{\tau}{e - 2\varepsilon}$, $m = \frac{\tau}{e - 2\varepsilon}$

Pullers: (O, A) and (T, R) Population shares: $o = \frac{\mu}{e - 2\varepsilon}$, $t = 1 - \frac{\mu}{e - 2\varepsilon}$

provided $\mu < \frac{\varepsilon(e - 2\varepsilon)}{e - \varepsilon}$ and $\tau < e - 2\varepsilon$

Table 1 presents a numerical example comparing economic performance in the monomorphic population and at the bimorphic focal point. In the monomorphic equilibrium all pushers offer the minimum and pullers always accept. There is no mindsight among pushers or transparency among pullers. All pullers are act-rational and all pushers believe that all pullers are act-rational. Mindsight, transparency and rule-rationality exist along closed orbits around the bimorphic focal point. In these populations too offers are never rejected since blind pushers believe that pullers are rule-rational and offer almost everything. The total product realized is only τ less than in the monomorphic equilibrium, but is allocated almost entirely to the pullers. Mindsight and transparency thus serve to reverse the allocation in favor of pullers.

<Table 1 about here>

Without mindsight pushers exploit the act-rational pullers and this is a stable equilibrium. Since mindsight and transparency enable pullers to turn the tables and exploit the pushers, it can be said that pullers prefer to display their rule-rational decision logic but pushers prefer not to look. Yet even though mindsight hurts them, pushers with mindsight can be present in an evolved population. Although mindsight hurts pushers, because it is locally adaptive it does not go extinct. As the numerical example shows, even a small fraction of pushers with mindsight may be enough to support transparency and rule-rationality among almost all pullers, and make the blind pushers adopt rule-rationality as their theory of the puller's mind.

According to experimental evidence compiled across numerous ultimatum experiments conducted in different cultures, on average, human proposers offer 40% of the pie and human responders reject 16% of offers. (Oosterbeek 2004) Although our simple model cannot explain this data, it points to the possibility that a more refined model of mindsight, transparency and evolutionary commitment to rule-rationality may be able to help account for this evidence.

5. Trust island

An island is populated by two species: getters and workers. Every minute a random getter obtains a resource of value $e > 0$. The getter can "invest" some portion $x \in [0, e]$ of the resource to be worked on by a random worker who is nearby. The worker's effort multiplies the value of the investment by a factor of $k > 1$. After finishing the work, the worker can pay back any amount $y \in [0, kx]$ to the getter. The resulting payoffs are $\pi_1(x, y) = e - x + y$ to the getter and $\pi_2(x, y) = kx - y$ to the worker. The endowments in each dyad are: $\pi_1^0 = e$ and $\pi_2^0 = 0$. Under ideal institutions that costlessly ensure maximal investment in all dyads, the product realized is $P = ke$. This is the first-best baseline. We want to know: What kind of rationality may evolve on this island?

The decision logic of an act-rational worker never returns anything to the getter:

$$A(x, e) = 0$$

The decision logic of a rule-rational worker minimally rewards those getters who invest everything and punishes all others:

$$R(x, e) = \begin{cases} e + \varepsilon \text{ } if \text{ } x = e \\ 0 \text{ } if \text{ } x < e \end{cases}$$

Getter strategies are: $x_A = 0$ and $x_R = e$. The payoffs under the various combinations of decision logics are:

$$\begin{array}{ll} \pi_1^{AA} = e & \pi_2^{AA} = 0 \\ \pi_1^{RR} = e + \varepsilon & \pi_2^{RR} = ke - e - \varepsilon \\ \pi_1^{AR} = e & \pi_2^{AR} = 0 \\ \pi_1^{RA} = 0 & \pi_2^{RA} = ke \end{array}$$

By Proposition 3, the blind / opaque / act-rational population in which all getters are (B, A) and all workers are (O, A) is evolutionarily stable.

By Proposition 6, the following population is an evolutionary focal point:

Getters: (B, R) and (M, Φ^M) Population shares: $b = 1 - \frac{e + \varepsilon + \tau}{ke}$, $m = \frac{e + \varepsilon + \tau}{ke}$

Workers: (O, A) and (T, R) Population shares: $o = \frac{\mu}{e}$, $t = 1 - \frac{\mu}{e}$

provided $\mu < \frac{e\varepsilon}{e+\varepsilon}$ and $\tau < e(k-1)-\varepsilon$

Table 2 gives a numerical example comparing economic performance in the monomorphic population and at the bimorphic focal point. In the monomorphic equilibrium getters do not invest anything and the workers earn nothing. In orbits around the bimorphic focal point, blind getters believe that workers are committed to repay with interest and invest everything. Some of them are betrayed by act-rational workers. But investment occurs in most dyads, the only exception being dyads in which a getter with mindsight is paired with an opaque act-rational worker. Mindsight and transparency serve to increase the average product but also allocate most of the gains to the workers. However, average fitness of both getters and workers is higher near the bimorphic focal point than in the monomorphic population.

<Table 2 about here>

This case shows that mindsight, transparency and rule-rationality can be critical for trust, can make all players better off on average, and may exist in evolved populations. Unlike in the Ultimatum Game, mindsight and transparency are incentive-compatible for all: it can be said workers want to show their decision logic and getters want to see it. But since mindsight is costly, a fraction of getters evolves to free-ride without mindsight. Such blind trusting getters in turn create a niche for opaque act-rational workers, who evolve to prey on them. However, as the numerical example in Table 2 shows, distrust, betrayal, opaqueness and act-rationality can all be very rare even if only a minority of the getters have mindsight.

6. Conclusion

We studied the decision logics and the capabilities for showing and observing them which may evolve among randomly paired agents subject to selection based on their performance in a strategic interaction. We found if mindsight and transparency are costly and optional, universal blindness, opaqueness, and act-rationality constitute the only evolutionarily stable equilibrium. We also found that blind leaders and opaque responders cannot go extinct and opaque responders must be act-rational. We showed that mindsight, transparency, and rule-rationality cannot be maintained in the entire population, nor can they be maintained in an asymptotically stable share of the population.

However, we also demonstrated that the evolution of mindsight, transparency, and rule-rationality cannot be ruled out. Mindsight, transparency, and rule-rationality may be found in a significant share of the population, a share that fluctuates along a closed orbit around a focal point. In such populations, blind leaders assume every responder is rule-rational and leaders with mindsight assume that an opaque responder is act-rational.

We applied our general findings to two specific contexts: Ultimatum Game and Trust Game. In both games, mindsight, transparency, and rule-rationality serve to allocate most of the surplus to responders. Given the zero-sum nature of the Ultimatum Game, mindsight does not engender new value and leaders are better off in the equilibrium without mindsight. But in the Trust Game, both leaders and responders earn more in populations with mindsight and mindsight, transparency, and rule-rationality are essential for enabling the investment that generates the new value.

On stability and complexity grounds, we conclude that act-rationality, blindness, and opaqueness are more likely to be found in a population of evolved agents than rule-rationality, mindsight, and transparency. However, we also conclude that the evolution of rule-rationality, mindsight, and transparency is possible. Committing to, displaying, observing and inferring preferences constitutes an important dimension along which strategically interacting agents may evolve. Further work on this dimension of evolution may shed light on the complex psychology and behavior revealed in game experiments.

References

Abhijit, B., Weibull, J. 1995. Evolutionary selection and rational behavior. in Kirman, A., Salmon, M., eds. Learning and Rationality in Economics. Oxford:Blackwell, 343-363

Robert A., 2008. Rule-Rationality versus Act-Rationality. Discussion Paper 497, Center for Rationality and Interactive Decision Theory, Hebrew University

Danielson, P., 1992. Artificial Morality: Virtuous Robots for Virtual Games. London: Routledge

Gauthier, D., 1986. Morals by Agreement. Oxford: Oxford University Press

Gintis, H., 2009. Game Theory Evolving. 2nd ed. Princeton: Princeton University Press

Guth, W., Kliemt, H., 1998. The indirect evolutionary approach: bridging the gap between rationality and adaptation. Rationality and Society 10(3), 377-399

Guth, W., Kliemt, H., 2000. Evolutionarily stable co-operative commitments. Theory and Decision 49, 197-221

Guth, W., Peleg, B., 2001. When will payoff maximization survive? An indirect evolutionary analysis. Journal of Evolutionary Economics 11, 479-499

Guth, W., Yaari, M., 1992. An evolutionary approach to explain reciprocal behavior in a simple strategic game. in Witt, U., ed., Explaining process and change – Approaches to evolutionary economics, Ann Arbor: University of Michigan Press

Heifetz, A., Shannon, C., Spiegel, Y., 2007. What to maximize if you must. Journal of Economic Theory 133, 31-57

Nesse, R., ed., 2001. Evolution and the Capacity for Commitment, New York: Russell Sage Foundation

Oosterbeek, H., Sloof, R., van de Kuilen, G., 2004. Cultural differences in ultimatum game experiments: evidence from a meta-analysis. Experimental Economics 7, 171–188

Schelling, T., 1960. The strategy of conflict. Cambridge: Harvard University Press

Siegel, D., 2001. Toward an interpersonal neurobiology of the developing mind: attachment relationships, 'mindsight,' and neural integration. Infant Mental Health Journal, 22(1-2), 67-94

Appendix

Definition. Decision logics Θ and Θ' are *different* (denoted by $\Theta \neq \Theta'$) if $\Theta(x,e) \neq \Theta'(x,e)$ except at a countable set of points (x',e') where $\Theta(x,e) = \Theta'(x',e')$

Proof of Proposition 1 Suppose that in the population of responders there are some opaque agents with act-rational decision logic *A* and some opaque agents with a different decision logic $Z \neq A$. Since a leader facing an opaque responder in a dyad cannot see whether the responder has A or Z, the leader's action is the same against either type of responder. By definition of act-rationality, replying to the leader's action using *A* yields the responder a higher payoff than replying using *Z*. Thus opaque Z-responders earn lower average fitness than opaque A-responders. Therefore the assumed population is not a fixed point. Next suppose that all opaque responders are act-rational. A mutant opaque responder with decision logic *Z* has a lower fitness and therefore cannot invade. ■

Proof of Proposition 2 When a leader with mindsight is paired with an opaque responder, he sees that the responder is opaque and uses some theory of mind. Suppose some leaders use *A* and some leaders use $Z \neq A$ as the theory of mind for opaque responders. Since, at a fixed point all opaque responders are act-rational (Proposition 1), the leader who uses A earns a higher payoff than the leader who uses Z. Thus the two types of leader cannot coexist at a fixed point. Furthermore, if all leaders with mindsight use *A* for opaque responders, a mutant leader using *Z* for opaque responders cannot invade.

When a leader with mindsight meets a transparent responder with decision logic Θ , he can use it as the theory of mind or may use a different theory $Z \neq \Theta$. Using Z yields the leader a lower payoff, and thus lower fitness, than using Θ. Thus the two types of leader cannot coexist at a fixed point. Furthermore, if all leaders with mindsight use Θ for transparent responders, a mutant leader using *Z* for transparent responders cannot invade.■

Proof of Proposition 3 Suppose all leaders are of type (B, A) and all responders are of type (O, A). Since all responders are opaque, a mutant leader with mindsight can get no information but would incur the cost of mindsight. Since all responders are act-rational, a mutant leader with a different theory of mind would earn less fitness. Thus mutants with mindsight or different theory of mind cannot invade the leader population. Since all leaders are blind, a mutant responder who is transparent cannot influence any leader's action, but would incur the cost of transparency. A mutant responder who is not act-rational would earn less fitness than an act-rational responder. Thus mutant responders who are transparent or have a different decision logic cannot invade the responder population. That other possible monomorphic populations are not fixed points is proved as Propositions 4 and 5. ■

Proof of Proposition 4 Suppose all leaders have mindsight. By Proposition 2, if this is a fixed point then all leaders have theory of mind Φ^M. From the definition of R it follows that, when playing against a population of (M, Φ^M) leaders, a (T, R) responder would earn higher fitness than a (T, Z) responder for any $Z \neq A$. Assumption 2 implies that a (T, R) responder would earn higher fitness than a (O, A) responder. By Proposition 1, an (O, A) responder would earn higher fitness than a (O, Z) responder for any $Z \neq A$. Thus, if all leaders have mindsight and the population is a fixed point, then all responders must be (T, R). A mutant (B, R) leader would earn the same payoff as a (M, Φ^M) leader but save the cost of mindsight. Thus a mutant blind leader can invade the leader population. ■

Proof of Proposition 5 Suppose there is a fixed point at which all responders are of type (T, Θ). A blind (B, Θ) leader would play the same as a (M, Φ^M) leader, but save the cost of mindsight. Thus the leader population must consist entirely (B, Θ) agents. A mutant (O, A) responder can earn more against such leaders than a (T, Θ) responder. Therefore a mutant opaque act-rational responder can invade. ■

Proof of Proposition 6 Consider a population of leaders consisting of (B, R) and (M, Φ^M) types and a population of responders consisting of (O, A) and (T, R) types. The system is an asymmetric evolutionary game analyzed by Gintis (2009, Sec. 12.17). We follow his approach to solve for the fixed point and ascertain its stability.

Let *b* be the share of blind leaders, *m* the share of leaders with mindsight, *o* the share of opaque responders, and *t* the share of transparent responders. The leader population is given by $\mathbf{p} = (b, m)$, where $b + m = 1$ and $b, m \in [0,1]$. The responder population is given by $\mathbf{q} = (o, t)$, where $o + t = 1$ and $o, t \in [0,1]$. The payoff matrices are:

Leaders: $$\mathbf{\Pi}_L = \begin{bmatrix} \pi_1^{RA} & \pi_1^{RR} \\ \pi_1^{AA} - \mu & \pi_1^{RR} - \mu \end{bmatrix}$$

Responders: $$\mathbf{\Pi}_R = \begin{bmatrix} \pi_2^{RA} & \pi_2^{RR} - \tau \\ \pi_2^{AA} & \pi_2^{RR} - \tau \end{bmatrix}$$

Adding a constant to each entry in a column of $\mathbf{\Pi}_L$ or in a row of $\mathbf{\Pi}_R$ does not affect the replicator dynamics. Therefore we can simplify the payoff matrices as follows:

$$\mathbf{\Pi}_L' = \begin{bmatrix} 0 & \mu \\ \pi_1^{AA} - \pi_1^{RA} - \mu & 0 \end{bmatrix}$$

$$\mathbf{\Pi}_R' = \begin{bmatrix} 0 & \pi_2^{RR} - \pi_2^{RA} - \tau \\ \pi_2^{AA} - \pi_2^{RR} + \tau & 0 \end{bmatrix}$$

Using the population share of blind leaders *b* and population share of opaque responders *o* as state variables, we can express the replicator equations of the two populations as follows:

$$\dot{b} = b(1-b)(\alpha - \gamma b)$$
$$\dot{o} = o(1-o)(\beta - \delta o)$$

where

$$\alpha = \mu > 0$$
$$\beta = \pi_2^{AA} - \pi_2^{RR} + \tau < 0$$
$$\gamma = \pi_1^{AA} - \pi_1^{RA}$$
$$\delta = \pi_2^{AA} - \pi_2^{RA}$$

The fixed point is given by:

$$b^* = \frac{\beta}{\delta} = \frac{\pi_2^{AA} - \pi_2^{RR} + \tau}{\pi_2^{AA} - \pi_2^{RA}} \quad , \quad o^* = \frac{\alpha}{\gamma} = \frac{\mu}{\pi_1^{AA} - \pi_1^{RA}}$$

and $m^* = 1 - b^*$ and $t^* = 1 - o^*$. Since α and β have opposite signs, this population is an evolutionary focal point surrounded by trajectories which are closed orbits such that the time frequencies of (b, o) along the orbits equal (b^*, o^*). (Gintis, 2009, Theorem 12.9)

Next, we need to establish that a third type of responder cannot invade the bimorphic responder population consisting of (O, A) and (T, R) types. By Proposition 1, a mutant of type (O, Z≠A) cannot invade. A mutant of type (T, A) also cannot invade because it is treated the same as (O, A) responder by both types of leader and therefore earns τ less fitness than (O, A) responder. Lastly, consider a mutant responder of type (T, Z) such that Z≠R and Z≠A. Since R is the decision logic that induces the leader with mindsight to take the action which lets the responder maximize its payoff, the mutant earns less than the incumbent (T, R) responder earns against (B, R) or (M, Φ^M) leader.

Finally, we need to establish that a third type of leader cannot invade the bimorphic leader population consisting of (B, R) and (M, Φ^M) types. By Proposition 2, a leader with mindsight but with a theory of mind different from Φ^M cannot invade. So we only need to consider a mutant blind leader with decision logic Z≠R. The expected fitness of such (B, Z) leader is $V_{BZ} = o\pi_1^{ZA} + t\pi_1^{ZR}$. The expected fitness of incumbent (M, Φ^M) is $V_M = o\pi_1^{AA} + t\pi_1^{RR} - \mu$. The mutant cannot invade if $V_{BZ} < V_M$, which reduces to

$$\mu < o(\pi_1^{AA} - \pi_1^{ZA}) + t(\pi_1^{RR} - \pi_1^{ZR}) .$$

By definition of A, $\pi_1^{AA} > \pi_1^{ZA}$. By definition of R, $\pi_1^{RR} > \pi_1^{ZR}$. Therefore, the upper bound on μ is positive and for any Z there exists a sufficiently small positive μ such that (B, Z) cannot invade. ■

Proof of Proposition 7 The proof is analogous to the proof of Proposition 6. The difference lies in the fitness earned by blind leaders and by responders paired with blind leaders. Consider a population of leaders consisting of (B, A) and (M, Φ^M) types and a population of responders consisting of (O, A) and (T, R) types. The key parameters of the replicator dynamic are:

$$\alpha = \pi_1^{AR} - \pi_1^{RR} + \mu < 0$$
$$\beta = \pi_2^{AA} - \pi_2^{RR} + \tau < 0$$
$$\gamma = \pi_1^{AR} - \pi_1^{RR}$$
$$\delta = \pi_2^{AR} - \pi_2^{RR}$$

The fixed point is given by:

$$b^* = \frac{\beta}{\delta} = \frac{\pi_2^{AA} - \pi_2^{RR} + \tau}{\pi_2^{AR} - \pi_2^{RR}} \quad , \quad o^* = \frac{\alpha}{\gamma} = \frac{\pi_1^{AR} - \pi_1^{RR} + \mu}{\pi_1^{AR} - \pi_1^{RR}}$$

Since α and β have the same sign, *(b*, o*)* is a saddle point and therefore unstable. (Gintis, 2009, Theorem 12.9) ■

Table 1 Economic performance in the Ultimatum game in the monomorphic equilibrium consisting solely of act-rational agents without mindsight or transparency, and at the bimorphic focal point where act- and rule-rational agents coexist.

Base game: Ultimatum		e=100, ε=5, μ=2, τ=1	
		Monomorphic	Bimorphic
Population			
Leaders (pushers)			
Blind (B, A)	a	1	0
Blind (B, R)	b	0	0.989
Mindsight (M, Φ^M)	m	0	0.011
Responders (pullers)			
Opaque act-rational (O, A)	o	1	0.022
Transparent rule-rational (T, R)	t	0	0.978
Performance			
Leader average fitness	V_L	95	5
Responder average fitness	V_R	5	94
Total product realized	$P=V_L+V_R$	100	99
First-best product possible	e	100	100
Fraction of first-best realized	P/e	1	0.99
Leader share of product	V_L/P	0.95	0.05
Responder share of product	V_R/P	0.05	0.95
Fraction of dyads with rejected offers	a t	0	0

Table 2 Economic performance in the Trust game in the monomorphic equilibrium consisting solely of act-rational agents without mindsight or transparency, and at the bimorphic focal point where act- and rule-rational agents coexist.

Base game: Trust		e=100, ε=5, μ=2, τ=1, k=5	
		Monomorphic	Bimorphic
Population			
Leaders (getters)			
Blind (B, A)	a	1	0
Blind (B, R)	b	0	0.788
Mindsight (M, Φ^M)	m	0	0.212
Responders (workers)			
Opaque act-rational (O, A)	o	1	0.02
Transparent rule-rational (T, R)	t	0	0.98
Performance			
Leader average fitness	V_L	100	102.9
Responder average fitness	V_R	0	394
Total product realized	$P=V_L+V_R$	100	496.9
First-best product possible	k e	500	500
Fraction of first-best realized	P/(ke)	0.2	0.9938
Leader share of product	V_L/P	1	0.21
Responder share of product	V_R/P	0	0.79
Fraction of dyads with reciprocated trust	(b+m)t	0	0.98
Fraction of dyads with distrust	a + m o	1	0.00424
Fraction of dyads with betrayal	b o	0	0.01576