--DRAFT--

Evolution of mindsight: transparent agents and the preference to look at them

Dimitry Rtischev

Faculty of Economics
Gakushuin University
Tokyo, Japan

March 19, 2012

****Early draft. Corrections & suggestions welcome.****

Abstract

We extend evolution of preferences theory by endogenizing mindsight, the ability to observe the decision logic that another agent uses to choose his action in a strategic interaction. The agents we study are randomly paired to play a sequential-move game and are subject to evolutionary selection based on their performance in the game. The agents may be opaque or transparent, blind or with mindsight. Agents with mindsight observe the decision logic of transparent agents. We argue that consistent with Aumann's distinction between act-rationality and rule-rationality, evolution selects opaque agents to be act-rational and transparent agents to be rule-rational. We find that in the unique evolutionary equilibrium all agents are blind, opaque and act-rational. However, we also find that there exists an evolutionary focal point surrounded by closed orbits along which rule-rational transparent agents and agents with mindsight are present in significant proportions. We apply the theory to Ultimatum and Trust games and find that evolved populations with mindsight can exhibit significantly different economic performance than populations without mindsight.

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

1. Introduction

Models of strategic interaction usually assume that agents obey the following two directives: (D1) always choose the action that yields oneself the highest payoff, and (D2) assume every other agent always chooses the action that yields him the highest payoff. In the context of the one-shot anonymous Ultimatum Game to divide a dollar, a responder who obeys D1 accepts any offer of one cent or more, and a proposer who obeys D1 and D2 offers one cent, thereby earning 99. A large body of experimental evidence shows 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 D1 and most human proposers not obey D1 and/or D2. One possible reason is that players are influenced by commitments that have been left out of the model.

That commitments can drastically change the course of strategic interactions has been widely appreciated since Schelling's *Strategy of Conflict*, in which he observed that "it may be perfectly rational ... to wish for the power to suspend certain rational capabilities in particular situations." (Schelling 1960, p. 18) Recently, Aumann (2008) emphasized the distinction between "rule-rationality," which relies on commitment to a rule of behavior, and "act-rationality," which optimizes without rules, and argued that evolution favors agents who are rule-rational.[1] Evolution of preferences theory has explored the evolution of agents committed to maximize a "subjective utility" which may differ from the objective payoffs. (Banerjee and Weilbull 1995; Guth and Peleg 2001) The theory established that in most games evolutionary selection produces agents committed to pursue subjective utility that does differ from objective payoffs. Asking "What to maximize if you must" in a generic game, Heifetz, *et. al.* (2007) formally demonstrated that strategic interaction inherently generates the incentive to commit to maximize something other than the objective payoffs, and such commitments do not disappear under evolutionary dynamics.

An essential assumption underlying these results is that agents can make credible commitments and induce other agents to take such commitments into account. To what extent this assumption holds in human interactions is a central issue in evolutionary psychology. To evolutionary psychologists, it is naïve to analyze the Ultimatum Game as if it were the whole game, because even anonymous strangers meeting to play only once are a product of a long process of evolution during which there was selection pressure to become committed and able to perceive or infer the commitments of others. Evolutionary psychology views human interactions as fundamentally mediated by theory of mind, by subjective commitments secured by emotions, and by other psychological capabilities shaped by evolution. (Nesse 2001) At the level of the brain, interpersonal neurobiology

[1] Using different nomenclature, the distinction between rule-rationality and act-rationality has also been made by philosophers of morality and rationality. (*e.g.*, Gauthier, 1986, Chapter VI)

studies "mindsight," the ability to form and make use of mental representations of how another human thinks and feels in a given interaction. (Siegel 2001)

One specific example of the evolved human capacity to show and observe mental states is gaze-following studied by Tomasello *et. al.* (2007). Summarizing evidence that "human eyes are colored in a way that helps advertise both their presence and their gaze direction much more saliently than in other primates" they cite studies showing that:

- In a sample of 81 species (including humans), 80 species were found to have low contrast in eye and facial skin coloration (i.e., the outline of the eyes and the position of the iris were difficult to distinguish due to the similarity in color of the facial skin, sclera, and iris). Humans were the only species in which the eye outline and the position of the iris were clearly visible, since the exposed sclera was paler than the lightest colored iris or surrounding skin.
- The human eye and its visible regions were found to be disproportionately large and horizontally elongated for body size *(i.e.*, the visible regions of human eyes were bigger than that of the much larger gorilla). The amount of visible sclera was three times greater in humans than in orangutans when looking straight ahead and twice as large when looking to the side.

Tomasello *et. al.* conducted experiments showing that the information made available by the human eye design is used by humans but not by apes. Thus, the human equilibrium is "disclose gaze direction to others / observe gaze direction of others" whereas the ape equilibrium is "do not disclose / do not observe." Arguably, showing and observing each other's gaze direction is just one small component of mindsight. But the fact that it has evolved in humans and not other primates lends support to the view that mindsight capabilities may help explain unique aspects of human interactions that are not found among other species or among the psychologically simplistic agents of economic theory.

Taking hints from evolutionary psychology, this paper attempts to extend evolution of preferences theory by endogenizing mindsight. The agents we study are randomly paired to play a sequential-move game and are subject to evolutionary selection based on their performance in the game. The agents may be opaque or transparent, blind or with mindsight. Only agents with mindsight observe the decision logic of transparent agents. The decision logic of opaque agents is not observable. Mindsight and transparency are costly.

We find that in the unique evolutionary equilibrium all agents are blind, opaque and act-rational. However, we also find that there exists an evolutionary focal point surrounded by closed orbits along which rule-rational transparent agents and agents with mindsight can be present in significant proportions. Although the evolutionary basis for blindness/opaqueness/act-rationality is much stronger than for mindsight/transparency/rule-rationality, the latter can exist in evolved populations and have a significant effect on behavior. We examine our results in two specific contexts – Ultimatum and Trust games -- and find that evolved populations with mindsight can exhibit significantly different economic performance than populations without mindsight.

Previous attempts in the evolution of preferences literature to model how agents get information about other agents' subjective utility or strategic commitments have mostly adopted the paradigm noisy signaling. (Guth and Kliemt, 1998 and 2000; Heifetz, *et. al.*, 2007) In such models, an agent observes the type of another agent but with exogenous noise that results in a positive probability of error. By endogenizing mindsight and transparency in a deterministic framework, our approach allows focus on the evolutionary aspects of the problem without the complications of stochastic issues.

Philosophers working on rationality and morality have gone further in explicitly considering mindsight and transparency among strategically interacting agents. Danielson's (1990) pioneering book, in which he algorithmically examines Gauthier's (1986) theory of rational morality, conceives agents as logic programs that may examine other agents' programs and may allow themselves to be examined by others' programs. Like Danielson, we let each agent operate according to its own decision logic and allow agents that show and observe these decision logics. Unlike Danielson, we assume mindsight and transparency are costly, develop a general framework with a generic base game, and analyze evolutionary population dynamics.

The rest of the paper is organized as follows. The next section lays out a formal framework for analyzing evolution of mindsight and Section 3 presents equilibrium analysis. We then apply the results to Ultimatum Game (Section 4) and 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 leader from the population of leaders and one responder the population of responders. A parameter e drawn from a given probability distribution describes the state of the environment (*e.g.,* weather) that prevails at the time the dyad interacts.[2] Each dyad plays a base game as follows: first both players observe e, then the leader 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 follower is endowed with π_2^0 . If either or both players in a dyad abstain from the game, both players keep these endowments.

[2] The assumption of the changing environment requires agents to dynamically compute their actions when they play rather than be hardwired to always play the same action.

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).

Definition An agent's *decision logic* is the function, algorithm, or program that determines how the agent chooses actions.

Definition The decision logic of an agent *with mindsight* takes as input the decision logic of a transparent agent.

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

Definition The decision logic of a *blind* agent cannot take as input the decision logic of any other agent.

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

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] $\mathbf{q} = (q_1,\ldots,q_O,q_{O+1},\ldots,q_{O+T})$, 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 $\sum q_i = 1$.

A leader is programmed to act according to decision logic Λ. The leader's action is given by $x = \Lambda(\Phi,e)$, where Φ is the decision logic of the responder as believed by the leader. In other words, Φ is the leader's theory of the responder's mind. Every leader maximizes its own payoff given his belief in how the responder will react; that is the decision logic of a leader with theory of mind Φ is

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

The type of leader is specified by (λ,Φ), where $\lambda \in \{B,M\}$ indicates blindness or mindsight. 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 $\mathbf{p} = (p_1,\ldots,p_B,p_{B+1},\ldots,p_{B+M})$, where $p_i \in [0,1]$ is the share of the i^{th}

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

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 $\sum p_i = 1$.

In terms of psychological traits, there are four possible types of dyads: MT, MO, BT, and BO. Mindsight operates only in the MT dyad: the leader's decision logic takes the responder's decision logic as input and computes the action $x = \Lambda(\Theta, e)$. In the MO, BT, and BO dyads, the leader cannot observe Θ and therefore the leader's decision logic relies on a built-in theory of mind to compute $x = \Lambda(\Phi, e)$. 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*)

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

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

Blind-Transparent dyad: (*i = 1,...,B*; *j = O+1,...,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,...,B+M*; *j = 1,...,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,...,B+M*; *j = O+1,...,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:

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

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

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 O+T rows and B+M 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: $\dot{p}_i = p_i(V_i^L - \overline{V}_L), \quad i = 1,\ldots,B+M$

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

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

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

3. Equilibrium analysis

Consider the following responder decision logics that correspond to Aumann's (2008) act-rationality and rule-rationality:

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

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

$\forall e \quad \nexists R' \ s.t. \ \pi_2(\Lambda(R'), R'(\Lambda(R')), e) > \pi_2(\Lambda(R), R(\Lambda(R)), e)$ 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.

We will denote the leader's payoff-maximizing strategy given the decision logic of the responder as follows:

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

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

We use the following shorthand notation to denote base game payoffs to leaders (*i=1*) and responders (*i=2*), net 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:

$$x_A \neq x_R \;\forall e$$

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

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

In order to make it possible for mindsight and transparency to evolve, it is necessary to assume that the cost of these capabilities is not too large relative to how they influence 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 identify which types of agents may exist in 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 all leaders with mindsight believe that opaque responders are 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 monomorphic populations of blind leaders (B, A) and opaque act-rational responders (O, A) constitute an evolutionarily stable fixed point.

Proposition 4 There does not exist a fixed point at which all leaders have mindsight.

Proposition 5 There does not exist a fixed point at which all responders are transparent and all have the same decision logic.

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 *(b*, o*)*.

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 which combinations of psychological traits and decision logics may evolve among agents subject to selection based on their performance in a sequential dyadic interaction. The only evolutionarily stable state is universal blindness, opaqueness, and act-rationality. Blind leaders and opaque responders cannot go extinct. Opaque responders must be act-rational. Although mindsight, transparency, and rule-rationality cannot be universal and cannot be present in asymptotically stable proportions, they may be present in populations along closed erdogic orbits around a fixed 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: leaders and responders. At every unit of time, a leader and a responder randomly meet near a resource of value *e*. Extracting the resource requires them to cooperate. The leader proposes to give *x* to the responder after they extract the resource, where $0 < \varepsilon <= x <= e - \varepsilon$ and ε is the minimum amount that can be allocated to an agent.[5] The responder accepts or rejects the offer. If the responder rejects, the agents go their separate ways and the resource rots away. If the responder 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 proposer and responder 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, the total average product realized on the island per unit of time is $P = e$. This is the first-best baseline.

The decision logic of an act-rational responder is:

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

The decision logic of a rule-rational responder is:

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

Proposer 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 leaders are (B, A) and all responders are (O, A) is evolutionarily stable.

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

[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.

Leaders: (B, R) and (M, Φ^M) Pop. shares: $b = 1 - \dfrac{\tau}{e-2\varepsilon}$, $m = \dfrac{\tau}{e-2\varepsilon}$

Responders: (O, A) and (T, R) Pop. shares: $o = \dfrac{\mu}{e-2\varepsilon}$, $t = 1 - \dfrac{\mu}{e-2\varepsilon}$

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

Table 1 presents a numerical example comparing economic performance in the monomorphic and bimorphic populations. In the monomorphic equilibrium all proposers offer the minimum and responders always accept. There is no mindsight among proposers or transparency among responders. All responders are act-rational and all proposers believe that all responders are act-rational. Mindsight, transparency and rule-rationality exist along closed orbits around the bimorphic fixed point. In these populations too offers are never rejected since blind proposers believe that responders 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 responders. Mindsight and transparency thus serve to reverse the allocation in favor of responders.

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

According to experimental evidence compiled across numerous studies 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 and evolutionary commitment to rule-rationality may be able to help account for this evidence.

Table 1 Economic performance in the ultimatum game in populations with and without mindsight

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

5. Trust island

An island is populated by two species: leaders and responders. At every unit of time, a leader and a responder meet near a resource of value $e > 0$. The leader can keep the entire resource to himself or can "invest" some portion $x \in [0,e]$ of it in a project managed by the responder. The responder works to multiply the value of the investment by a factor of $k > 1$. The responder can then pay back any amount $y \in [0,kx]$ to the leader. The resulting payoffs are $\pi_1(x,y) = e - x + y$ to the leader and $\pi_2(x,y) = kx - y$ to the responder. The endowments that the agents in each dyad keep if one or both abstain from playing are: $\pi_1^0 = e$ and $\pi_2^0 = 0$. First-best institutions that costlessly ensure maximal investment in all dyads generate a total average product of $P = ke$. This is the first-best baseline.

The decision logic of an act-rational responder never returns anything to the leader:

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

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

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

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

$$\begin{array}{ll} \pi_1^{AA}=e & \pi_2^{AA}=0 \\ \pi_1^{RR}=e+\varepsilon & \pi_2^{RR}=ke-e-\varepsilon \\ \pi_1^{AR}=e & \pi_0^{AR}=0 \\ \pi_1^{RA}=0 & \pi_2^{RA}=ke \end{array}$$

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

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

Leaders: (B, R) and (M, Φ^{M}) Pop. shares: $b = 1-\dfrac{e+\varepsilon+\tau}{ke}$, $m=\dfrac{e+\varepsilon+\tau}{ke}$

Responders: (O, A) and (T, R) Pop. shares: $o=\dfrac{\mu}{e}$, $t=1-\dfrac{\mu}{e}$

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

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

This case shows that mindsight, transparency and rule-rationality can be critical for trust, can make all players better off, and may exist in evolved populations. Unlike in the Ultimatum Game, mindsight and transparency are incentive-compatible for all: responders want to show their decision logic and leaders want to see it. But since mindsight is costly, a fraction of leaders evolve to free-ride without mindsight. Such blind trusting leaders in turn create a niche for opaque act-rational responders, 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 leaders have mindsight.

Table 2 Economic performance in the trust game in populations with and without mindsight

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

6. Conclusion

We identified the decision logics and mindsight-related capabilities that may evolve among randomly paired agents subject to selection based on their performance in a sequential interaction. We found that the state of universal blindness, opaqueness, and act-rationality is the unique evolutionarily stable equilibrium. Blind leaders and opaque responders cannot go extinct and opaque responders evolve to be act-rational. Mindsight, transparency, and rule-rationality cannot be universal and cannot be present in asymptotically stable proportions. However, mindsight, transparency, and rule-rationality are not ruled out by evolution. Mindsight, transparency, and rule-rationality may comprise significant share of populations along closed erdogic orbits around a fixed point. In such populations, blind leaders assume each responder is rule-rational and leaders with mindsight assume that an opaque responder is act-rational.

We examined our findings in 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 critical for investment that generates new surplus.

Both on stability and complexity grounds, we conclude that act-rationality, blindness, and opaqueness are more likely to be found in evolved populations than rule-rationality, mindsight, and transparency. But although we found that populations with rule-rationality, mindsight, and transparency are at best fluctuating around a fixed point, we also conclude that mindsight may constitute an important dimension along which agents with complex psychologies evolve and may be critical for understanding behavior among such agents.

References

Abhijit Banerjee and J. Weilbull (1995) “Evolutionary selection and rational behavior” in A. Kirman and M. Salmon, eds., “Learning and Rationality in Economics” Blackwell, Oxford, 343-363

Robert Aumann (2008) “Rule-Rationality versus Act-Rationality” Discussion Paper 497, Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem

Peter Danielson (1992) *Artificial Morality: Virtuous Robots for Virtual Games*, Routledge, London

David Gauthier (1986) *Morals by Agreement*, Oxford University Press, Oxford

Herbert Gintis (2009) *Game Theory Evolving*, 2nd ed., Princeton University Press, Princeton

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

Werner Guth and H. Kliemt (2000) "Evolutionarily stable co-operative commitments" *Theory and Decision* 49, 197-221

Werner Guth and B. Peleg (2001) "When will payoff maximization survive? An indirect evolutionary analysis" *Journal of Evolutionary Economics* 11, 479-499

Aviad Heifetz, C. Shannon, and Y. Spiegel (2007) "What to maximize if you must" *Journal of Economic Theory* 133, 31-57

Randolph M. Nesse, ed. (2001) *Evolution and the Capacity for Commitment*, Russell Sage Foundation, New York

Hessel Oosterbeek, R. Sloof, and G. van de Kuilen (2004) "Cultural Differences in Ultimatum Game Experiments: evidence from a Meta-Analysis" *Experimental Economics* 7:171–188

Thomas C. Schelling (1960) *The Strategy of Conflict*, Harvard University, Cambridge

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

Michael Tomasello, B. Hare, H. Lehmann , J. Call (2007) "Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis" *Journal of Human Evolution* 52:3, 314-320

Appendix

Model in the case of bimorphic population

The model introduced in Section 2 allows for an arbitrary number of decision logics. In the special case of a bimorphic population we use the following simplified notation.

Population state

Leaders: (B, R) and (M, Φ^M) types only. $\mathbf{p} = (b, m)$, $b + m = 1$

Responders: (O, A) and (T, R) types only. $\mathbf{q} = (o, t)$, $o + t = 1$

The payoff matrices summarizing fitness earned in each of the four dyads are respectively:

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

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

Expected fitness of the two leader types

$$V_{BR}(o,t) = o\pi_1^{RA} + t\pi_1^{RR}$$
$$V_M(o,t) = o\pi_1^{AA} + t\pi_1^{RR} - \mu$$

Expected fitness of the two responder types

$$V_{OA}(b,m) = m\pi_2^{AA} + b\pi_2^{RA}$$
$$V_{TR}(b,m) = \pi_2^{RR} - \tau$$

Population average fitness:

Leaders: $\bar{V}_L \equiv bV_{BR} + mV_M$

Responders: $\bar{V}_R \equiv oV_{OA} + tV_{TR}$

Replicator dynamic:

Leaders: $\dot{b} = b(V_{BR} - \bar{V}_L)$

Responders: $\dot{o} = o(V_{OA} - \bar{V}_R)$

Proof of Proposition 1 Suppose there is a type of opaque responder in a fixed point population that has decision logic A and another type with a different decision logic Z ≠ A. Since each type of responder is opaque, every type of leader plays the same action against each. By definition of act-rationality, replying to the leader's action using A yields a higher payoff than using any other decision logic Z. Thus opaque responders who use Z have lower average fitness than act-rational responders, which implies the population is not a fixed point. It also follows that if all responders are opaque and act-rational, a mutant opaque responder with decision logic Z cannot invade. ■

Proof of Proposition 2 By Proposition 1, at a fixed point all opaque responders are act-rational. A leader with mindsight who believes an opaque responder is not act-rational earns lower average fitness than a leader with mindsight who believes an opaque responder is act-rational. Therefore, leaders with mindsight who believe an opaque responder is not act-rational cannot coexist at a fixed point with leaders with mindsight who believe an opaque responder is act-rational. It also follows that if all leaders with mindsight believe that opaque responders are act-rational, a mutant leader with different theory of mind 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. ■

Proof of Proposition 4 Suppose there is a fixed point at which all leaders have mindsight. Assumption 2 implies that all responders must be transparent and rule-rational, since any other type of responder would earn lower average fitness. A blind leader who believes responders are rule-rational would earn the same payoff in the base game as a leader with mindsight but save the cost of mindsight. Thus a mutant leader of type (B, R) 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 leader with theory of mind $\Phi = \Theta$ would earn the same payoff in the base game as a leader with mindsight but save the cost of mindsight. Thus the leader population must consist entirely of agents of type (B, Θ). But an opaque act-rational responder (O, A) can earn more against such leaders than a transparent responder (T, Θ). Therefore a mutant responder of type (O, A) can invade if the responder population consists entirely of (T, Θ) responders. ■

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.

Adding a constant to each entry in a column of $\mathbf{\Pi}_L$ or in a row of $\mathbf{\Pi}_F$ 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. Consider a mutant leader of type (B, A).

Its expected fitness is $V_{BA} = o\pi_1^{AA} + t\pi_1^{AR}$. 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 $V_M > V_{BA}$, which reduces to $\mu < (\pi_1^{RR} - \pi_1^{AR})t$. At the fixed point t*, this condition is

$$\mu < \frac{(\pi_1^{AA} - \pi_1^{RA})(\pi_1^{RR} - \pi_1^{AR})}{(\pi_1^{AA} - \pi_1^{RA}) + (\pi_1^{RR} - \pi_1^{AR})}$$

Lastly, consider a mutant leader of type (B, Z) such that Z≠R and Z≠A. Such a mutant earns less against (O, A) responder than (B, A) leader and less against (T, R) responder than (B, R) leader. ■

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:

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

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) ■