THE ROYAL SOCIETY

Published online 17 February 2003

Dressing the mind properly for the game

David Sally

Cornell University, Johnson Graduate School of Management, 371 Sage Hall, Ithaca, NY 14853 6201, USA
(*dfs12@cornell.edu*)

Game theory as a theoretical and empirical approach to interaction has spread from economics to psychology, political science, sociology and biology. Numerous social interactions—foraging, talking, trusting, coordinating, competing—can be formally represented in a game with specific rules and strategies. These same interactions seem to rely on an interweaving of mental selves, but an effective strategy need not depend on explicit strategizing and higher mental capabilities, as less sentient creatures or even lines of software can play similar games. Human players are distinct because we are less consistent and our choices respond to elements of the setting that appear to be strategically insignificant. Recent analyses of this variable response have yielded a number of insights into the mental approach of human players: we often mentalize, but not always; we are endowed with social preferences; we distinguish among various types of opponents; we manifest different personalities; we are often guided by security concerns; and our strategic sophistication is usually modest.

Keywords: game theory; theory of mind; rationality; social preferences; risk dominance; strategic sophistication

1. INTRODUCTION

> Society is an interweaving and interworking of mental selves. I imagine your mind, and especially what your mind thinks about my mind, and what your mind thinks about what my mind thinks about your mind. I dress my mind before yours and expect that you will dress yours before mine. Whoever cannot or will not perform these feats is not properly in the game.
>
> (Cooley 1927, pp. 200–201)

If society is an interweaving of mental selves, then games are a particularly useful way to look more closely at the quality and hang of the fabric. To play a game is to engage in a certain kind of interaction, and the general claim of the research has been that the way players dress their minds here is indicative of their mental attire in other social situations. Games have become a critical theoretical and empirical tool in the social and biological sciences. A game is a formal representation of an interaction among strategies and 'strategy-carriers'. Examples of such carriers in the game theory literature range from automata, lines of software, viruses, hungry vampire bats, protective fishes, colluding corporations and warring countries, to individual human beings. An explosion of experimental work in the past 20 years has shown that this last category of strategy-carriers, despite their advantages in the areas of reasoning, rationality and mentalizing, can be the most befuddling and the least consistent game-players. These assays have shown humans to be at various times cooperative, altruistic, competitive, selfish, generous, equitable, spiteful, communicative, distant, similar, mindreading or mindblind as small elements in the game structure or social setting are altered.

There is no possible way to do justice to the wealth of work, the thousands of studies and the manifold models that comprise modern game theory in this paper. Rather, I will focus on recent results from behavioural economics and social and cognitive psychology that detail some of the mental apparel donned by human players under various social conditions. Just as we might dress unthinkingly and automatically in the early morning hours and knowingly grab a light jacket on a windy afternoon, just as we might wear socks every day and a suit on only special occasions, there are similar, discernible patterns in the mental garments we display while playing different games. Human players are not the buttoned-up, conservative, uniform dressers that early game theory expected them to be. Rather, our mental dress is much more casual, simple and flexible. The key findings of our mental fashion review are the following:

(i) Mentalizing is employed in many, but not all, games, as are rules and norms.
(ii) The way the game is displayed significantly affects strategies.
(iii) Players bring strong social preferences to a game.
(iv) Games can be used to diagnose individual differences and personality consistencies.
(v) Concerns about risk and security can determine player choices.
(vi) In general, most players are not very strategically sophisticated.

One contribution of 15 to a Theme Issue 'Decoding, imitating and influencing the actions of others: the mechanisms of social interaction'.

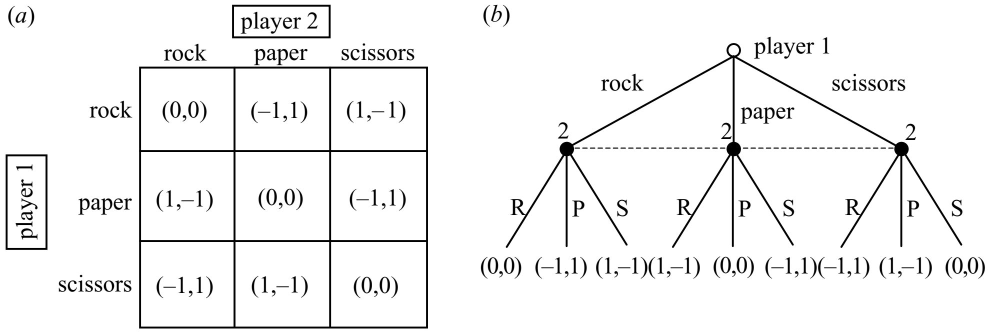

Figure 1. The rock, paper, scissors game. (*a*) Normal form, (*b*) extensive form.

2. FORMAL WEAR: MORNING SUITS AND DINNER JACKETS

As the inventors of game theory stated, formally a game 'is simply the totality of the rules which describe it' (Von Neumann & Morgenstern 1944). The key rules are those that govern the number of players, their possible moves, the flow of information and the outcomes resulting from terminal combinations of moves. A strategy is an action plan developed by a player that uses all the information available and that prescribes a move at each stage of the game. Outcomes are usually represented by simple payoff functions. In the standard theory, these are utility functions of the form

$$v_i = u_i(\pi_i(a_i,a_{-i})), \tag{2.1}$$

where $\pi_i(a_i,a_{-i})$ is the payoff to player i resulting from his/her own action and those of all other players (a_{-i}). The standard restrictions placed on the function, $u_i(\cdot)$, allow analysis to be based on a direct identification of utility, v_i, with the payoff.

We can distinguish some important categories of games. Zero-sum games are those in which any gain for a single player is offset by an equivalent loss spread across all other players; all other games present players with the opportunity to create or destroy common value. Games can be repeated or single-shot. Finally, they can be normal form or extensive, that is, drawn as a matrix or a tree. The simple zero-sum children's game of rock, paper, scissors is displayed in both normal and extensive forms in figure 1. The payoffs here are normalized so that the winner and loser receive 1 and –1, respectively, and a draw gives each player 0. The dotted line, an information set, connecting the second-stage nodes in figure 1*b* indicates that player 2 is uncertain which of the branches he/she is on as he/she chooses. Information sets allow any game to be displayed as either a box or a tree, a cross-dressing that proves theoretically that player strategies should be independent of the game form.

One solution concept that is both prescriptive and descriptive is the Nash equilibrium. An equilibrium is a combination of player strategies that is stable. Stability in a Nash equilibrium arises because each player's strategy is a best response to the strategies of all of the other players (Nash 1950). To see this, imagine that, before a single move was made, each player announced honestly to all of the other players what he or she was tentatively planning to do. These tentative plans form a Nash equilibrium if each player answers 'no' to the following question: 'Given what everyone else just said, is there a strategy that would make me better off than my tentative plan does?' Here, suppose that our children announced 'rock' and 'scissors'. This is not a Nash equilibrium because the scissors player wants to change to 'paper'. There are, in fact, no pairs of strategies that would be confirmed; the only truthful dialogue that would cause no subsequent change is if each says: 'I have no idea what I'm going to do' or 'I could choose anything'. Hence, the only Nash equilibrium has each child randomizing uniformly among her three moves. If the game were repeated, this equilibrium would result in each player winning, losing and drawing one-third of the games. These were exactly the proportions found in an imaging study of this game conducted by Gallagher *et al.* (2002).

The Nash equilibrium can involve a back-and-forth mentalizing process that was most vividly dramatized in an encounter in the film, *The Princess Bride*. The game here involves a duel over two goblets of wine, one of which has seemingly been poisoned by the mysterious Man in Black. Vizzini must choose one of the goblets and attempts to read the cloaked knowledge and masked intentions of the Man in Black as follows.

> But it's so simple. All I have to do is divine from what I know of you. Are you the sort of man who would put the poison into his own goblet, or his enemy's? … Now, a clever man would put the poison into his own goblet, because he would know that only a great fool would reach for what he was given. I'm not a great fool, so I can clearly not choose the wine in front of you. But you must have known I was not a great fool; you would have counted on it, so I can clearly not choose the wine in front of me … You've beaten my giant, which means you're exceptionally strong. So, you could have put the poison in your own goblet, trusting on your strength to save you. So I can clearly not choose the wine in front of you. But, you've also bested my Spaniard which means you must have studied. And in studying, you must have learned that man is mortal so you would have put the poison as far from yourself as possible, so I can clearly not choose the wine in front of me.
>
> (Goldman 2003, p. 2)

Vizzini finally brings his mentalizing to a close, drinks and dies, not realizing that both goblets were poisoned, the Man in Black having built an immunity to the toxin.

The plight of this sophisticated rationalizer was due to his ignorance of the other's resistance. In general, however, Aumann & Brandenburger (1995) showed that if two players' actions are based on mutual knowledge of

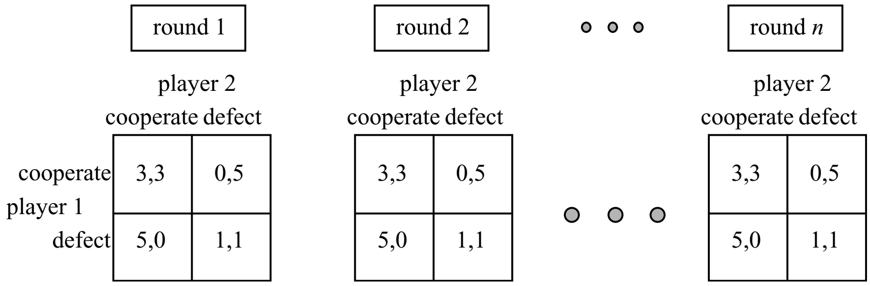

Figure 2. A finitely repeated Prisoner's Dilemma.

their payoff functions, of their rationality and of their strategic conjectures, then these actions will constitute a Nash equilibrium. Knowledge of another's conjectures can be derived through the application of theory of mind—our awareness, as vividly portrayed by Vizzini, that others have mental states that differ from our own and that this unique cerebral dress explains their behaviours (Frith & Frith 1999). Hence, the formal result above means that for a two-person game a first-order theory of mind is sufficient to support a Nash equilibrium: a player does not have to interpret what the other's conjectures of his/her conjectures are.

Demands on mentalizing increase when there are more than two players. It can be proved formally that with three or more players in a game, common knowledge of the others' strategic conjectures is now required to support a rational, premeditated Nash equilibrium (Aumann & Brandenburger 1995). Common knowledge of intentions requires higher orders of theory of mind to perceive levels of conjectures about conjectures. Also, many of the most frequently studied games become much more difficult in practice as the roster of players increases. For those researchers exploring the relationship between mentalizing and games, these facts indicate that experiments involving more than two players may be worth pursuing for they can sometimes be an even more acute test of theory of mind.

Cooley might call Vizzini and any player who hews to the Nash equilibrium in every game, who is supremely logical and self-interested, a rational dandy, one whose mind is clothed in the most sophisticated morning coat or evening wear. Such a person might find himself in a finitely repeated Prisoner's Dilemma as shown in figure 2. The Prisoner's Dilemma has been used in hundreds of experiments and models to portray the conflict between doing what is best for the individual (defection) and helping maximize the group's outcome (cooperation) (for comprehensive reviews see Sally (1995) and Allison *et al.* (1996)). The only Nash equilibrium is to defect in each of the n rounds, as any thoughts of cooperation are banished by the inevitability of defection. For example, the popinjay might hypothesize a completely cooperative relationship up through the last round, but then he/she would see that there is an incentive to defect in the nth trial. Since the rational counterpart knows this as well, mutual defection is assured, making the cooperative relationship last for $n - 1$ rounds. With another cycle of theory of mind, he/she realizes that both he/she and the other will recognize the $n - 1$ round as the new 'last' round and will defect here as well, cutting mutual cooperation down to $n - 2$ rounds. This inevitable logic, usually called backward induction, unzips the cooperative relationship completely.

There is only one problem with this story: backward induction is rarely, if ever, seen among real players (Johnson *et al.* 2002). Most players will cooperate for many of the rounds, not defecting until the last or second to last trial, if they defect at all (data: Sally 1995; theory: Kreps *et al.* 1982). This is but one example of a general finding: most players are not rational dandies and most games do not encourage or reward such naive and calculating behaviour. In fact, the formal definition of a game given above misstates its reality: 'The game, one would like to say, has not only rules but also a *point*' (Wittgenstein 1958, § 564). Much of the recent research on game theory has been directed towards determining what these points might be and in what ways people will regularly vary from the Nash equilibrium.

3. EMPEROR'S NEW CLOTHES?

The emperor was convinced as he marched through court to his throne that his robes were luxurious, his breeches well fitted and his blouses radiant. Of course, well suited, or dressed at all, he was not. In the same way, *contra* Cooley, perhaps we believe that our minds are dressed when we participate in society, but really they are quite naked. Perhaps mentalizing is a grand illusion.

We have learned from studies of those with autistic spectrum disorder that mental attire is not incidental to the interweaving of larger society. Many researchers agree that one of the core deficits of autism is a damaged theory of mind (Baron-Cohen *et al.* 1985, 1993; Frith *et al.* 1991). An autistic individual manifests this deficit in the laboratory in an inability to interpret stories involving false knowledge, bluffs and deceptions (Happé 1994) and to automatically ascribe motivations, intentions and emotions to moving and interacting abstract figures (Castelli *et al.* 2002; by contrast, see Heider & Simmel 1944). Largely due to their malfunctioning theory of mind, those with autism, in keeping with Cooley's statement, are not properly in many critical social games (Sally 2001). For example, they infrequently engage in spontaneous pretend play (Carruthers 1996); their language tends to be overly literal and devoid of metaphor, irony, implication and indirectness (Mitchell *et al.* 1997; Tager-Flusberg 2000); they find fictional drama to be unrewarding or frustrating (Sacks 1995); their friendship and acquaintance networks are sparsely peopled (Frith *et al.* 1994); and their gifts are inclined to be incongruous (Park 1998).

Autism is an organic brain disorder with multiple (known and unidentified) causes that injure the innate theory of mind and hamper its full development. Scanning studies of those with autism and those with focal brain lesions indicate that the theory of mind is found in a distributed neural system incorporating the medial prefrontal cortex, which includes areas activated in monitoring the self's inner states, and the superior temporal sulcus, which is associated with the detection of the movement of animate objects, especially eyes, hands and mouth (Sabbagh & Taylor 2000; Frith & Frith 2000; Puce & Perrett 2003; Frith & Frith 2003). So, this neural system may be the physical loom upon which the interworking of mental selves occurs and the social games of friendship, language, gift-giving, etc., are woven. The activity of specific neurons indicates that mentalizing is not a majestic illusion in these games.

Still, it could be true that in the formal games of concern here, as the social trappings are removed to reveal the bare bones of the underlying matrix or branching tree, the minds of players also are unclothed despite their sensations of strategic finery. It is possible that it is not the game itself but its social setting that promotes mentalizing, and in the starkness of the laboratory our minds are simple and unadorned. There are two very recent sources of evidence relevant here as follows:

(i) imaging studies that examine the brains of active participants in games; and
(ii) a study comparing the decisions of autistic and control subjects.

Gallagher *et al.* (2002) had normal subjects play the rock, paper, scissors game of figure 1 while positron emission tomography was employed to document their neural activity. In two conditions subjects were told that they were playing either another person or a rule-following, pre-programmed computer. In addition, those facing the first type of opponent were explicitly encouraged to outwit and outguess him/her, while half of those facing the computer were told to just randomize each round across the moves. The one region of significant difference in the brain activity of these experimental groups appeared in the most anterior portion of the paracingulate cortex bilaterally—a region solidly within the hypothesized theory-of-mind neural system (Frith & Frith 2003). There is an another implication of this study: there are approaches to games in which full mentalizing does not occur. Facing a computer and just generating random moves does not engage the medial prefrontal cortex in the same way that outsmarting a human counterpart does.

A similar comparison, but a different game, was used by Rilling *et al.* (2002). These researchers scanned the brains of players matched with another person or a computer in a finitely repeated Prisoner's Dilemma (figure 2). The scanning took place both when the players were deciding on a move and after an outcome was revealed. Relative to the other three outcomes, mutual cooperation with a human partner elicited more activity in a variety of areas of the brain associated with reward processing—the medial prefrontal cortex (Brodmann's area 11) and rostral anterior cingulate gyrus (BA 32). By contrast to the previous study, there was overlap in activation between human and machine: the ventromedial/orbitofrontal cortex also responded to mutual cooperation with the computer. As there were no special instructions regarding the computer opponent, these experimental participants may have anthropomorphically and quite naturally mentalized. Finally, the decision to reciprocate the human counterpart's cooperative move in the prior round also evoked BA 32, whose role may be to bring an emotional tone to the theory-of-mind system (Bush *et al.* 2000; Frith & Frith 2003).

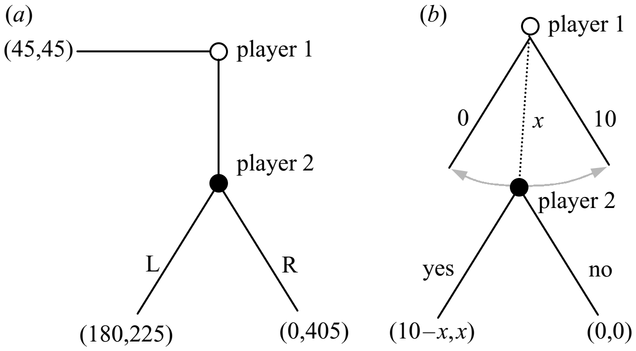

Figure 3. Two extensive-form bargaining games. (*a*) Trust game, (*b*) ultimatum game.

The extensive-form game shown in figure 3*a* is representative of those used by a number of researchers. In this 'trust game,' player 1 must decide whether to end the game right away with matching payoffs of 45 for each player, or to pass the decision-making power to player 2 who must then choose to move left or right, resulting in outcome pairs of (180, 225) or (0, 405), respectively. A decision to continue the game implies that the first player trusts the second to not be overly greedy and choose the fairer outcome. Rigdon (2002) compared choices made in trust games that were presented in the theoretically equivalent normal and extensive forms, and found that the tree form increased the frequency of the most cooperative outcome from 15% to 50%. In an imaging study (McCabe *et al.* 2001), participants played extensive-form trust games against human and computer opponents, and they were told all the details of the latter's probabilistic strategy. A human–computer comparison of brain activation in those subjects who cooperated at least one-third of the time overall showed heightened responses in a number of areas including the medial prefrontal cortex.

These imaging studies indicate that whether a game evokes mentalizing can depend on the identity of the opponent, the form of the game, and whether the strategic approach is one of problem solving in the form of rule detection and rule application. E. Hill and D. Sally (unpublished data) reported somewhat similar findings from their research on autism and games. Among children who played a finitely repeated Prisoner's Dilemma (figure 2), the degree of development of theory of mind was correlated with greater levels of cooperation. However, among adults, those with autism and those normal adults with less sophisticated and accurate mentalizing skills were less reactive to changes in the settings of the games. Theory of mind seemed to be used by the control adults to compete more vigorously and cooperate more thoroughly, in accordance with the altered rules of the game.

Autistic and control subjects also played the ultimatum game in figure 3*b*. This game represents the essence of many bargaining situations. The first player is given 10 points and must make an offer, x, of some portion of the total to the second player. The latter can say 'yes' or 'no', resulting in payoffs of $(10 - x, x)$ and (0,0), respectively. In other words, the second player can scotch the whole deal if x is too small or unfair. There were the many interesting comparisons and contrasts among the participants, adults and children, controls and autistics, including the following: if a child failed a second-order false belief test (which was much more probable among the autistics), then he/she was significantly more prone to offer nothing or only one point to the second player. Alternatively, most of the children who passed this test initially offered an even split of the total points. Most autistic adults proposed offers of either zero or 50%, while adults without autism tendered bids that were strategically shaded a notch or two below half.

These and other distinctions between the autistics and the controls, though significant, were outnumbered by the similarities. Hill & Sally (2002) claimed that this was a surprising finding since the literatures of behavioural economics, developmental psychology and cognitive neuroscience predicted gross differences. The decisions of autistic participants were not founded on a well-developed, innate theory of mind but, rather, on a rule-based, awkward, compensatory mechanism (Happé 1994), and yet, they were often reasonable approximations of those of the mentalizers. This type of norm-based behaviour has been witnessed in other games (Henrich *et al.* 2001), and it raises a question for future researchers—what is the relationship between theory of mind and norms? Is one a necessary input to the other? In what social games are they complementary or discordant?

Taken as a group, these studies provide compelling testimony that mentalizing does occur in formal games: the emperor is at least partially clad some of the time. It may be that the extensive form encourages theory of mind, but we do not know why. Does the tree illustrate movement and progress; does it reduce uncertainty; does it underscore intentions, and if so, how? Mentalizing does not always occur naturally. Gallagher *et al.* (2002) felt they had to advocate for it in their game, and Hoffman *et al.* (2000) raised the average offer in their ultimatum game by recommending that subjects consider the other's expectations. To the questions about norms raised above, one might ask, what is the relationship between learning within a repeated game and mentalizing? Lastly, what elements of the identity of the counterpart make the neurally based theory of mind wax and wane?

4. MAO JACKETS AND PINSTRIPED SUITS

One reason that strategy with a machine can be quite distinct is that it is difficult for a player to identify with his/her silicon-based counterpart. Clothing, whether it is donned for a party rally in the Forbidden City or a meeting in the board room, is often used to express solidarity, and it should come as no surprise that mental dress frequently fulfils the same function. Researchers have determined that most players bring social preferences to a game, preferences that place some weight on the intentions and outcomes of other players.

A variety of functional forms have been proposed to represent these social preferences (Edgeworth 1881; Sawyer 1966; Loewenstein *et al.* 1989; Rabin 1993; Montgomery 1994; Levine 1998; Fehr & Schmidt 1999; Sethi & Somanathan 2001; Charness & Rabin 2002). These proposals place a non-zero weight on the other players' payoffs (unlike the asocial utility of equation (2.1)), and most include factors representing pure altruism and reciprocity. Sally (2000) reviewed the history of thought on social preferences within a variety of disciplines and developed a modern theory of sympathy. This model of sympathy is a recasting of the original theory developed by Hume (1740), Smith (1790) and Darwin (1936), and it is a variant of these recent approaches—the balance theory of Heider (1958), the simulation theory of Goldman (1989), the notion of identification of Coleman (1990) and the concept of empathy of Gallese (2003). Following the tradition of functional forms listed here and borrowing the mathematics of kinship altruism as developed by Hamilton (1964) and Hirshleifer (1978), one version of sympathetic preferences (for a two-person game) is the following:

$$v_i = \pi_i + (1 - \alpha_i)\exp(-\alpha_i\varphi_{ij}\psi_{ij})\pi_j. \quad (4.1)$$

The pure altruism of player i is α_i, and φ_{ij} and ψ_{ij} are player i's perceived physical and psychological distances from player j. Simply, the closer another is to us, the more readily we have fellow-feeling, identification, sympathy for them.

Personal consistency (α_i) across a variety of games is the focus of the burgeoning literature on social values orientation. Individuals are categorized as one of three types—prosocial, individualistic, competitive—corresponding to $\alpha_i < 1$, $\alpha_i = 1$, $\alpha_i > 1$, respectively. A panel of decomposed games, i.e. sets of paired payoffs without any moves, is used to screen and identify a person's orientation (Messick & McClintock 1968). Arithmetically, a smaller α_i means a heavier weight on the other's payoff and a greater return from an other-oriented strategy in a particular game. Hence, a prosocial orientation fosters cooperation in the Prisoner's Dilemma (McClintock & Liebrand 1988), helping behaviour (McClintock & Allison 1989) and more productive negotiations (DeDreu *et al.* 2000). In a very interesting study on the developmental aspects of social value orientation, Van Lange *et al.* (1997) discovered that individuals with a prosocial orientation had more siblings and more sisters than did the other two types. Also, across the lifespan the proportion of individualists and competitors decreases from 45% in the 15–29 age group to 18% in the over 60 group. As yet, there is no research addressing the neural foundations of α_i or how it might be related to theory of mind.

Equation (4.1) predicts that physical and psychological closeness will also motivate prosocial individuals to employ an other-oriented strategy in a given game. Indeed, physical proximity will promote cooperation in the repeated Prisoner's Dilemma and generosity in bargaining games (Wichman 1970; Michelini 1971; Bohnet & Frey 1999). In addition, psychological similarity and familiarity will support prosocial behaviour in these same games (McNeel & Reid 1975; Hoffman *et al.* 1996). Finally, in

(*a*)

		player 2 left	player 2 right
player 1	left	1,1	0,0
	right	0,0	1,1

(*b*)

		player 2 S	player 2 D
player 1	S	5,5	0,4
	D	4,0	2,2

Figure 4. Two coordination games. (*a*) Pure coordination, (*b*) impure coordination.

a protocol that allowed participants to meet each other, identify commonalities, and then play a 'trust' game, participants were significantly more likely to trust than were the anonymous, distant subjects in previous experiments (Glaeser *et al.* 2000).

The cognitive assessment of distance is inextricably bound with the emotions (Hume 1740; Zajonc 1998), in particular with how much we like another person and how good we believe another to be (Heider 1958). So, perceptions of distance interact with evaluation and attraction in such social phenomena as clustered friendship networks in schools and workplaces (Newcomb 1956; Segal 1974), an instinctive approach towards the good and avoidance of the bad (Solarz 1960; Bargh 1997), a positive evaluation of something solely because it is near to us (Cacioppo *et al.* 1993) and affection for someone who mimics our gestures (Chartrand & Bargh 1998). (For a comprehensive review of this evidence on sympathy, see Sally (2000).)

It is a simple and yet profound point that most social interaction involves physical closeness. The result, as Goffman (1983) points out, is that 'emotion, mood, cognition, bodily orientation, and muscular effort are intrinsically involved, introducing an inevitable psychobiological element' (p. 5). There is no doubt that this psychobiological element has a real impact on the strategies players employ in a variety of games. Emerging work on the innateness, pervasiveness and mechanisms of imitation (Meltzoff & Decety 2003), mirror neurons, simulation, and sympathy (Gallese 2003), and the perception of motion, intentions and goals (Csibra 2003; Johnson 2003) promises to generate insight into player strategizing and social preferences. This research might find game theory an attractive arena in which to rigorously test hypotheses about interaction.

5. CHAIN MAIL AND MACS

The primary purpose of some habiliments is to protect the body beneath: for armour, from an aggressive opponent; for raincoats, from the elements. When they invented game theory, Von Neumann & Morgenstern (1944) analysed many games from a 'maximin' perspective. This solution focuses on what move in a game would maximize the player's worst, or minimum, payoff. A conservative, risk-averse player might play maximin in order to preclude large losses and boost the guaranteed outcome.

Similar security concerns lie behind the solution concept of risk dominance. To explore this principle, we need to examine another category of interaction—coordination games. Two examples are shown in figure 4. The first (figure 4*a*) is a problem of pure coordination: players need to pick the same move and mismatching is costly to both; however, no equilibrium is better than another. This game can represent an encounter between two people moving in opposite directions who will pass successfully only if they both move either to their left or their right (Schelling 1960), or the semantics shared by a speaker and a listener, i.e. 'left' and 'right' meaning left and right instead of one of the other infinite possibilities (Lewis 1969). Convention is often used to solve these problems: on the roadways, depending on the country, all drivers use the left or the right lane. Difficulties arise in less defined spaces such as rectangular mall parking lots where the drivers slow down dramatically, reducing simultaneity of decision, transforming the game into its extensive form equivalent, and allowing them to pass each other on the 'wrong' side if necessary. (The same extensive, sequential, unconventional solution can happen on the sidewalk as pedestrians are able to identify a potential miscoordination and take unilateral action to prevent it.) In the semantic game, convention is essential and is documented in the dictionary. Finally, players able to solve the semantic game will be able to communicate and untangle other coordination problems more easily, especially if only one person can speak (e.g. 'you go left') or they can talk sequentially (Cooper *et al.* 1992).

The 'impure' coordination game shown in figure 4*b* has two Nash equilibria—(S, S) and (D, D). The first equilibrium is payoff dominant relative to the second as both players have a higher return. However, the second is risk dominant as it arises from a move that is psychologically more secure and is confirmed by a wider range of beliefs about the other's move (it also, in this case, happens to be maximin as well). More generally, with two Nash equilibria, (a'_1, a'_2) and (a_1^*, a_2^*), the first is strictly payoff dominant if

$$\pi_1(a'_1,a'_2) > \pi_1(a_1^*,a_2^*), \quad \pi_2(a'_1,a'_2) > \pi_2(a_1^*,a_2^*) \tag{5.1}$$

and the second is risk dominant if

$$(\pi_1(a_1^*,a_2^*) - \pi_1(a'_1,a_2^*))(\pi_2(a_1^*,a_2^*) - \pi_2(a_1^*,a'_2)) > (\pi_1(a'_1,a'_2) - \pi_1(a_1^*,a'_2))(\pi_2(a'_1,a'_2) - \pi_2(a'_1,a_2^*)). \tag{5.2}$$

Although it seems obvious that payoff dominance should trump risk dominance (and does so when one equilibrium's return is far greater than the rest), game theorists have discovered that many evolutionary models and experimental participants are guided by risk dominance. In a simulated ecology of matrix games, reproductive success accrued to players who favoured the risk-dominant action (Kandori *et al.* 1993; Young 1993). Subjects in numerous tests of matrices like the one in figure 4*b* were very likely to coordinate on the risk-dominant (D, D) cell (Cooper *et al.* 1992; Straub 1995; Battalio *et al.* 2001; Schmidt *et al.* 2003).

So, risk dominance is a central principle of choice, but are there any essential impure coordination games? One example is that of the stag hunt in which the success of a group of hunters closing in on their prey depends on the absence of any gaps in the narrowing circle. If a single hunter falters, the stag will escape (Battalio *et al.* 2001). Such a weakest-link structure can be found in many team or group production situations, whether on the factory floor, the executive suite or the game-show stage.

Another instance may be the utterance game studied by pragmatics (Sally 2002, 2003). The game in figure 4*b* is created by the first player saying to the second, 'you are such a moron'. This utterance has a multiplicity of meanings and possible implications: the speaker could be intending to denigrate or praise the listener's intelligence. This particular array of payoffs can be shown to arise from a combination of distinctive cognitive effects (e.g. emphasis, surprise, negative politeness), altered processing costs and the efficiency loss of miscoordination (see Sally (2002) for details). The literal interpretation (Dumb, Dumb) risk dominates the ironic interpretation (Smart, Smart), while the latter is payoff dominant.

A number of authors have asserted that meaning coordination is determined by maximizing the difference between the benefits and costs of interpretation (Prieto 1966; Bourdieu 1977; Parikh 1991). The most important of these efforts is the principle of optimal relevance (Sperber & Wilson 1995).

(i) The utterance is relevant enough for it to be worth the addressee's effort to process it.
(ii) The utterance is the most relevant one compatible with the communicator's abilities and preferences.

The first part of the principle guarantees the addressee a positive payoff, and the second part that the outcome will be payoff dominant. However, as the research concludes, payoff dominance is no guarantee of selection. The addressee has security concerns as well and his/her mind may be draped with protective garb.

There can be situations, then, in which optimal relevance is trumped by risk dominance. Empirically, the interpretation of 'you are such a moron' depends on the relationship between the speaker and listener, not just their individual abilities and efforts: friends will coordinate on (S, S) while acquaintances or strangers will settle on (D, D). In experimental settings, subjects were more likely to interpret insult or sarcasm non-literally and to view it as appropriate verbal behaviour if the phrase was spoken between friends (Slugoski & Turnbull 1988; Jorgensen 1996; Kreuz *et al.* 1999). Technically, this arises when the social preferences of equation (4.1) are inserted into the calculation of risk dominance in equation (5.2) and, if the players are physically and psychologically close enough, the payoff-dominant cell becomes also risk dominant (Sally 2002). There is, then, a connection between sympathy, mindreading and identification and perceptions of risk dominance.

The impure coordination game of utterance meaning may be neurally represented in the right hemisphere of the brain. Individuals who have sustained damage to this hemisphere, as well as those with autism, have problems picking out punch lines to jokes, revising initial interpretations, understanding metaphors and comprehending discourse (Van Lancker 1997). The summary by Brownell *et al.* (2000) of the neurological evidence incorporates the model of a deciding player and an uncertain matrix.

> The right hemisphere is more likely than the left to process weak or diffuse associations and low frequency alternative meanings, and the right hemisphere maintains activation over longer prime-target intervals. The left hemisphere actively dampens or inhibits activation of alternatives and focuses on the contextually most dominant reading.
>
> (Brownell *et al.* 2000, p. 320)

For those researchers studying connections between the mind and brain, coordination games offer the chance to examine more closely a number of elements of social interaction: norm- and rule-based behaviours, risk dominance and the effect of uncertainty, the role of the language coordination game in other social games.

6. FASHION ON THE CUTTING EDGE

View the plight of the one who would be a fashion leader: this person has to be able to predict the taste of the 'hoi polloi' and then stay just one step ahead. It is a cutting edge: half a step ahead and the leader is undistinguished, two steps ahead and the leader is bizarre. The quandary of the stylish dresser is the same as that of the equities and bonds trader: ideally, you are one step ahead of the average dealer—half a step and transaction costs eat up your gains; two steps and either the rest of the world never makes it out to you or holding costs erode profits while you wait. Keynes (1936) originally drew this comparison between investment in stock and the selection of something (or someone) stylish or beautiful.

> It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, not even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be.
>
> (Keynes 1936, p. 156)

A new set of experiments has emerged recently in game theory to test what degree of intelligence people really employ when trying to guess what the majority will do. These games and protocols may be of interest to those studying cognition from a developmental or neural perspective.

Nagel (1995) staged a version of the beauty contest of Keynes (1936). Instead of a stylish outfit, 15–18 (N) participants had to choose an integer, x_i, between 0 and 100. The choices would be averaged and then the average would be multiplied by a factor, p. In one condition, for example, $p = 0.5$. A prize was offered to the individual whose guess was closest to the target, $p(\Sigma x_i/N)$. The only Nash equilibrium, in the example, has every player guessing zero. To see the logic of this, suppose the mean guess is greater than zero. Then at least one player is above the mean and that player, holding everyone else's choice fixed, can be made better off by reducing his/her number. As this is true for every player, the guesses should unravel all the way down the number line to zero.

As with the other games we have considered, the Nash equilibrium is a very unlikely outcome in this game unless players are experienced. Rather, one can distinguish the degrees of strategic intelligence proposed by Keynes. Zero-step players decide randomly; one-step players make the best response to opponents who are all zero step; two-step players make the best response to opponents who are all one step; etc. If the zero-step players randomize across the entire number range of 100, then the expected mean of their choices is 50. Accordingly, one-step players will target 25 (shaded by a little to account for their own impact on the overall mean), and two-step players will point towards 12. Nagel (1995) discovered that most of her subjects were either one- or two-step players; a finding replicated by Ho *et al.* (1995) in a greater variety of 'beauty contest' games.

A single or double layer of strategic sophistication was also the dress chosen by most of the participants in a study by Costa-Gomes *et al.* (2001). A wide range of two-person, normal-form games confronted each subject. The twist in the protocol was that all the payoffs displayed on a computer screen were masked. To learn what their own or the other's payoff was in a given cell, the subject had to move the cursor over the appropriate region in the matrix. The subject's information search could be recorded and analysed. (MouseLab, the name of this laboratory technology, is the equivalent of a mind's-eye-movement tracking device. It was developed by Payne *et al.* (1993) and has been applied in decision-making experiments by Camerer *et al.* (1993), X. Gabaix and D. Laibson (unpublished data), and others.)

A one-step player in a normal-form game just has to determine their own payoffs as they are assuming that their counterpart is randomizing across the available moves. A two-step player has to search across all the outcomes for both participants and needs to make certain paired comparisons. By combining search information with actual choices, Costa-Gomes *et al.* (2001) estimated that 45% of their subjects were one-step and 44% were two-step players.

There is clearly a great deal of overlap between the concepts of mentalizing and strategic sophistication. One might be cut from the whole cloth of the other. If they are not identical, the latter process may be more manifest in the anonymous, multi-player setting epitomized by the beauty contest, and the other may be more important in the more intimate repeated dyadic game. There may also be differences with respect to the search for information and the effects of uncertainty. Those scientists researching theory of mind may be able to help the game theorists get a better feel for the texture and grain of strategic sophistication.

7. HEADWEAR: CAPS AND HATS

For a number of decades, the formal models and empirical tests of game theory developed in parallel with relatively little contact. The tweed jacket of the theorist and the laboratory coat of the experimenter were rarely observed together. Recent developments in the field, only some of which have been reviewed here, have attempted to knit the models and observed decisions into a single fabric. New games have been invented; old games have been modified; new technologies, such as imaging and MouseLab, have been applied; game forms, trees and boxes, have been distinguished; risk dominance and social preferences have been appreciated; new pools of players, autistics and members of small-scale societies, have been analysed. In all, the varied mental garments of players have been more completely identified.

Cognitive psychologists, neuroscientists and animal behaviourists can contribute mightily to a greater understanding of decisions and actions within games. In turn, the uniting of theory and data in modern game theory may make its models and procedures more valuable for these scientists. The time is ripe for cross-disciplinary collaboration. Here are some ideas:

(i) Can we adapt the protocols discussed above to diagnose the strategic sophistication of chimpanzees or children? What is its developmental path? Similar questions could be asked with respect to risk dominance.
(ii) Imaging studies have isolated the neural system that controls the value of a perceived object (Tremblay & Schultz 1999). Scanning and MouseLab could be used together as people play a masked game. Would the value-related neural system be activated by the revelation of a large own payoff? What areas of the brain are stimulated when the payoffs of the counterpart are unmasked?
(iii) How are a prosocial or competitive orientation and, more generally, sympathy and theory of mind connected? Also, how are rule-based behaviour and mentalizing distinguished?
(iv) The extensive form may promote the use of theory of mind in a game. What other changes in the game structure or environment would produce the level of mentalizing we believe exists in a casual conversation? Or, contrary to our current belief, are many social games, including conversation, conducted with undressed minds?

Lastly, even the essay is an example of the interweaving of mental selves that Cooley postulated. As occurs in all social interactions, such as a sales call, the essayist dresses his mind for the reader as he writes. With a nod to Tversky (1977) one might ask, how is the essayist like a well-dressed salesperson? He dons his best suit, starts with a greeting and a little small talk, makes his pitch, remains ever polite, crafts his offer, asks for an agreement and ends with a tip of his cap.

REFERENCES

Allison, S. T., Beggan, J. K. & Midgley, E. H. 1996 The quest for 'similar instances' and 'simultaneous possibilities': metaphors in social dilemma research. *J. Personality Social Psychol.* **71**, 479–497.

Aumann, R. & Brandenburger, A. 1995 Epistemic conditions for Nash equilibrium. *Econometrica* **63**, 1161–1180.

Bargh, J. A. 1997 The automaticity of everyday life. *Adv. Social Cogn* **10**, 1–61.

Baron-Cohen, S., Leslie, A. & Frith, U. 1985 Does the autistic child have a 'theory of mind'? *Cognition* **21**, 37–46.

Baron-Cohen, S., Tager-Flusberg, H. & Cohen, D. J. 1993 *Understanding other minds: perspectives from autism*. Oxford University Press.

Battalio, R., Samuelson, L. & Van Huyck, J. 2001 Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica* **69**, 749–764.

Bohnet, I. & Frey, B. S. 1999 The sound of silence in prisoners' dilemma and dictator games. *J. Econ. Behav. Org.* **38**, 43–57.

Bourdieu, P. 1977 The economics of linguistic exchanges. *Social Sci. Inform.* **16**, 645–668.

Brownell, H., Griffin, R., Winner, E., Friedman, O. & Happé, F. 2000 Cerebral lateralization and theory of mind. In *Understanding other minds: perspectives from developmental cognitive neuroscience*, 2nd edn (ed. S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen), pp. 306–333. Oxford University Press.

Bush, G., Luu, P. & Posner, M. I. 2000 Cognition and emotional infuences in anterior cingulate cortex. *Trends Cogn. Sci.* **4**, 215–222.

Cacioppo, J. T., Priester, J. R. & Berntson, G. G. 1993 Rudimentary determinants of attitudes. II. Arm flexion and extension have differential effects on attitudes. *J. Personality Social Psychol.* **65**, 5–17.

Camerer, C., Johnson, E., Rymon, T. & Sen, S. 1993 Cognition and framing in sequential bargaining for gains and losses. In *Frontiers of game theory* (ed. K. Binmore, A. Kirman & P. Tani), pp. 27–47. Cambridge, MA: MIT Press.

Carruthers, P. 1996 Autism as mind-blindness: an elaboration and partial defence. In *Theories of theories of mind* (ed. P. Carruthers & P. K. Smith), pp. 257–274. Cambridge University Press.

Castelli, F., Frith, C., Happé, F. & Frith, U. 2002 Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* **125**, 1839–1849.

Charness, G. & Rabin, M. 2002 Understanding social preferences with simple tests. *Q. J. Econ.* **117**, 817–869.

Chartrand, T. L. & Bargh, J. A. 1998 The chameleon effect: the perception-behavior link and social interaction. *J. Personality Social Psychol.* **76**, 893–910.

Coleman, J. S. 1990 *Foundations of social theory*. Cambridge, MA: Belknap.

Cooley, C. H. 1927 *Life and the student*. New York: Knopf.

Cooper, R., DeJong, D. V., Forsythe, R. & Ross, T. W. 1992 Communication in coordination games. *Q. J. Econ.* **107**, 739–771.

Costa-Gomes, M., Crawford, V. P. & Broseta, B. 2001 Cognition and behavior in normal-form games: an experimental study. *Econometrica* **69**, 1193–1235.

Csibra, G. 2003 Teleological and referential understanding of action in infancy. *Phil. Trans. R. Soc. Lond.* B **358**, 447–458. (DOI 10.1098/rstb.2002.1235.)

Darwin, C. 1936 *The origin of the species and the descent of man*. New York: The Modern Library.

De Dreu, K. W., Weingart, L. R. & Kwon, S. 2000 Influence of social motives on integrative negotiations: a meta-analytic review and test of two theories. *J. Personality Social Psychol.* **78**, 889–905.

Edgeworth, F. Y. 1881 *Mathematical psychics*. New York: Augustus M. Kelley (reprinted 1967).

Fehr, E. & Schmidt, K. M. 1999 A theory of fairness, competition, and cooperation. *Q. J. Econ.* **114**, 817–868.

Frith, C. D. & Frith, U. 1999 Interacting minds—a biological basis. *Science* **286**, 1692–1695.

Frith, C. D. & Frith, U. 2000 The physiological basis of theory of mind: functional neuroimaging studies. In *Understanding other minds: perspectives from developmental cognitive neuroscience*, 2nd edn (ed. S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen), pp. 334–356. Oxford University Press.

Frith, U. & Frith, C. D. 2003 Development and neurophysiology of mentalizing. *Phil. Trans. R. Soc. Lond.* B **358**, 459–473. (DOI 10.1098/rstb.2002.1218.)

Frith, U., Morton, J. & Leslie, A. M. 1991 The cognitive basis of a biological disorder: autism. *Trends Neurosci.* **14**, 433–438.

Frith, U., Happé, F. & Siddons, F. 1994 Autism and theory of mind in everyday life. *Social Dev.* **3**, 108–124.

Gallagher, H. L., Jack, A. I., Roepstorff, A. & Frith, C. D. 2002 Imaging the intentional stance in a competitive game. *NeuroImage* **16**, 814–821.

Gallese, V. 2003 The manifold nature of interpersonal relations: the quest for a common mechanism. *Phil. Trans. R. Soc. Lond.* B **358**, 517–528. (DOI 10.1098/rstb.2002.1234.)

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A. & Soutter, C. L. 2000 Measuring trust. *Q. J. Econ.* **115**, 811–846.

Goffman, E. 1983 The interaction order. *Am. Sociol. Rev.* **48**, 1–17.

Goldman, A. I. 1989 Interpretation psychologized. *Mind Lang.* **4**, 161–185.

Goldman, W. 2003 *The princess bride*. See http://www.heatherbryson.com/princess/pbscript.htm

Hamilton, W. D. 1964 The genetical evolution of social behavior: I and II. *J. Theor. Biol.* **7**, 1–52.

Happé, F. G. E. 1994 An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped and normal children and adults. *J. Autism Dev. Disorders* **24**, 129–154.

Heider, F. 1958 *The psychology of interpersonal relationships*. Hillsdale, HJ: Lawrence Erlbaum Associates.

Heider, F. & Simmel, M. 1944 An experimental study of apparent behavior. *Am. J. Psychol.* **57**, 243–259.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. & McElreath, R. 2001 In search of homo economicus: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* **91**, 73–78.

Hill, E. & Sally, D. 2002 *Dilemmas and bargains: theory-of-mind, cooperation and fairness*. Institute for Cognitive Neuroscience, University College of London.

Hirshleifer, J. 1978 Competition, cooperation, and conflict in economics and biology. *Am. Econ. Rev. Proc.* **68**, 238–243.

Ho, T. H., Camerer, C. & Weigelt, K. 1995 Iterated dominance and iterated best response in experimental 'p-beauty contests'. *Am. Econ. Rev.* **88**, 947–969.

Hoffman, E., McCabe, K. & Smith, V. L. 1996 Social distance and other-regarding behavior in dictator games. *Am. Econ. Rev.* **86**, 653–660.

Hoffman, E., McCabe, K. & Smith, V. L. 2000 The impact of exchange context on the activation of equity in ultimatum games. *Exp. Econ.* **3**, 5–9.

Hume, D. 1740 *A treatise of human nature*. Oxford University Press (reprinted 1978).

Johnson, S. C. 2003 Detecting agents. *Phil. Trans. R. Soc. Lond.* B **358**, 549–559. (DOI 10.1098/rstb.2002.1237.)

Johnson, E. J., Camerer, C., Sen, S. & Rymon, T. 2002 Detecting failures of backward induction: monitoring information search in sequential bargaining. *J. Econ. Theory* **104**, 16–47.

Jorgensen, J. 1996 The functions of sarcastic irony in speech. *J. Pragmatics* **26**, 613–634.

Kandori, M., Mailath, G. J. & Rob, R. 1993 Learning, mutation, and long run equilibria in games. *Econometrica* **61**, 29–56.

Keynes, J. M. 1936 *The genera theory of employment, interest and money*. New York: Harcourt Brace (reprinted 1964).

Kreps, D. M., Milgrom, P., Roberts, J. & Wilson, R. 1982 Rational cooperation in the finitely repeated prisoners' dilemma. *J. Econ. Theory* **27**, 245–252.

Kreuz, R. J., Kassler, M. A., Coppenrath, L. & Allen, B. M. 1999 Tag questions and common ground effects in the perception of verbal irony. *J. Pragmatics* **31**, 1685–1700.

Levine, D. K. 1998 Modeling altruism and spitefulness in experiments. *Rev. Econ. Dynam.* **1**, 593–622.

Lewis, D. K. 1969 *Convention*. Harvard University Press.

Loewenstein, G. F., Thompson, L. & Bazerman, M. H. 1989 Social utility and decision making in interpersonal contexts. *J. Personality Social Psychol.* **57**, 426–441.

McCabe, K., Houser, D., Ryan, L., Smith, V. & Trouard, T. 2001 A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* **98**, 11 832–11 835.

McClintock, C. G. & Allison, S. 1989 Social value orientation and helping behavior. *J. Appl. Social Psychol.* **19**, 353–362.

McClintock, C. G. & Liebrand, W. B. G. 1988 Role of interdependence structure, individual value orientation, and another's strategy in social decision making: a transformational analysis. *J. Personality Social Psychol.* **55**, 396–409.

McNeel, S. P. & Reid, E. C. 1975 Attitude similarity, social goals, and cooperation. *J. Conflict Resolution* **19**, 665–681.

Meltzoff, A. N. & Decety, J. 2003 What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Phil. Trans. R. Soc. Lond.* B **358**, 491–500. (DOI 10.1098/rstb.2002.1261.)

Messick, D. M. & McClintock, C. G. 1968 Motivational basis of choice in experimental games. *J. Exp. Social Psychol.* **4**, 1–25.

Michelini, R. L. 1971 Effects of prior interaction, contact, strategy, and expectation of meeting on game behavior and sentiment. *J. Conflict Resolution* **15**, 97–103.

Mitchell, P., Saltmarsh, R. & Russell, H. 1997 Overly literal interpretations of speech in autism: understanding that messages arise from minds. *J. Child Psychol. Psychiat.* **38**, 685–691.

Montgomery, J. 1994 Revisiting Tally's Corner: mainstream norms, cognitive dissonance, and underclass behavior. *Rationality Soc.* **6**, 462–488.

Nagel, R. 1995 Unraveling in guessing games: an experimental study. *Am. Econ. Rev.* **85**, 1313–1326.

Nash, J. 1950 Equilibrium points in n-person games. *Proc. Natl Acad. Sci. USA* **36**, 48–49.

Newcomb, T. M. 1956 The prediction of interpersonal attraction. *Am. Psychol.* **11**, 575–586.

Parikh, P. 1991 Communication and strategic inference. *Linguistics Phil.* **14**, 473–514.

Park, C. C. 1998 Exiting nirvana. *Am. Scholar* **67**(Spring), 28–42.

Payne, J., Bettman, J. & Johnson, E. 1993 *The adaptive decision maker*. Cambridge University Press.

Puce, A. & Perrett, D. 2003 Electrophysiology and brain imaging of biological motion. *Phil. Trans. R. Soc. Lond.* B **358**, 435–445. (DOI 10.1098/rstb.2002.1221.)

Prieto, L. J. 1966 *Messages et signaux*. Paris: P.U.F.

Rabin, M. 1993 Incorporating fairness into game theory and economics. *Am. Econ. Rev.* **83**, 1281–1302.

Rigdon, M. L. 2002 Theory of mind in two-person experiments. In *Encyclopedia of cognitive science* (ed. L. Nadel). London: Macmillan.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S. & Kilts, C. D. 2002 A neural basis for social cooperation. *Neuron* **35**, 395–405.

Sabbagh, M. A. & Taylor, M. 2000 Neural correlates of theory-of-mind reasoning: an event-related potential study. *Psychol. Sci.* **11**, 46–50.

Sacks, O. 1995 *An anthropologist on Mars*. New York: Knopf.

Sally, D. 1995 Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992. *Rationality Soc.* **7**, 58–92.

Sally, D. 2000 A general theory of sympathy, mind-reading, and social interaction, with an application to the prisoners' dilemma. *Social Sci. Inform.* **39**, 567–634.

Sally, D. 2001 Into the looking glass: discerning the social mind through the mindblind. *Adv. Group Processes* **18**, 99–128.

Sally, D. 2002 'What an ugly baby!' Risk dominance, sympathy, and the coordination of meaning. *Rationality Soc.* **14**, 78–108.

Sally, D. 2003 Risky speech: behavioral game theory and pragmatics. *J. Pragmatics* (Submitted.)

Sawyer, J. 1966 The altruism scale: a measure of co-operative, individualistic, and competitive interpersonal orientation. *Am. J. Sociol.* **71**, 407–416.

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

Schmidt, D., Shupp, R., Walker, J. M. & Ostrom, E. 2003 Playing safe in coordination games: the role of risk dominance, payoff dominance, social history, and reputation. *Games Econ. Behav.* (Submitted.)

Segal, M. W. 1974 Alphabet and attraction: an unobtrusive measure of the effect of propinquity in a field setting. *J. Personality Social Psychol.* **30**, 654–657.

Sethi, R. & Somanathan, E. 2001 Preference evolution and reciprocity. *J. Econ. Theory* **97**, 273–297.

Slugoski, B. R. & Turnbull, W. 1988 Cruel to be kind and kind to be cruel: sarcasm, banter and social relations. *J. Lang. Social Psychol.* **7**, 101–121.

Smith, A. 1790 *The theory of moral sentiments*. Indianapolis: Liberty Classics (reprinted 1976).

Solarz, A. 1960 Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *J. Exp. Psychol.* **59**, 239–245.

Sperber, D. & Wilson, D. 1995 *Relevance*, 2nd edn. Oxford: Blackwell.

Straub, P. G. 1995 Risk dominance and coordination failures in static games. *Q. Rev. Econ. Finance* **35**, 339–363.

Tager-Flusberg, H. 2000 Language and understanding minds: connections in autism. In *Understanding other minds: perspectives from developmental cognitive neuroscience*, 2nd edn (ed. S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen), pp. 124–149. Oxford University Press.

Tremblay, L. & Schultz, W. 1999 Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708.

Tversky, A. 1977 Features of similarity. *Psychol. Rev.* **84**, 327–352.

Van Lancker, D. 1997 Rags to riches: our increasing appreciation of cognitive and communicative abilities of the human right cerebral hemisphere. *Brain Lang.* **57**, 1–11.

Van Lange, P. A. M., Otten, W., De Bruin, E. M. N. & Joireman, J. A. 1997 Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *J. Personality Social Psychol.* **73**, 733–746.

Von Neumann, J. & Morgenstern, O. 1944 *Theory of games and economic behavior*. Princeton University Press.

Wichman, H. 1970 Effects of isolation and communication on cooperation in a two-person game. *J. Personality Social Psychol.* **16**, 114–120.

Wittgenstein, L. 1958 *Philosophical investigations* (tr. G. E. M. Anscombe). Oxford: Basil Blackwell.

Young, H. P. 1993 The evolution of conventions. *Econometrica* **61**, 57–84.

Zajonc, R. B. 1998 Emotions. In *The handbook of social psychology*, 4th edn (ed. D. T. Gilbert, S. T. Fiske & G. Lindzey), pp. 591–632. Boston, MA: McGraw-Hill.