Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Economic Psychology 27 (2006) 73–97
JOURNAL OF Economic Psychology
www.elsevier.com/locate/joep

The development of interpersonal strategy: Autism, theory-of-mind, cooperation and fairness

David Sally [a,*], Elisabeth Hill [b,1]

[a] Tuck School of Business at Dartmouth College, 100 Tuck Hall, Hanover, NH 03755, USA
[b] Department of Psychology, Goldsmiths College, University of London, New Cross, London SE14 6NW, United Kingdom

Available online 6 September 2005

Abstract

Mentalising is assumed to be involved in decision-making that is necessary to social interaction. We investigated the relationship between mentalising and three types of strategic games – Prisoners' Dilemma, Dictator and Ultimatum – in children with and without autistic spectrum disorders. Overall, the results revealed less dramatic differences than expected among the normally developing age groups and the children with autism, suggesting that in these laboratory tasks, mentalising skills are not always necessary. There were, nonetheless, some important findings. Young children were more cautious about initiating cooperation than their older peers and, in bargaining situations, they were less generous in their opening unilateral grants and over-solicitous of an empowered receiver. Participants with autism did have a harder time shifting strategy between versions of the Prisoners' Dilemma, and they were much more likely to accept low initial offers in the Ultimatum game and to refuse fair proposals. In addition, participants' measured mentalising abilities explain intentional and strategic behaviour within the prisoners' dilemma and the avoidance of unsuccessful ultimatum proposals.

Keywords: Autism; Bargaining; Cooperation; Dilemmas; Theory-of-mind

* Corresponding author. Tel.: +1 607 257 5758.
E-mail addresses: david.f.sally@dartmouth.edu (D. Sally), e.hill@gold.ac.uk (E. Hill).
1 Tel.: +44 020 7919 7886; fax: +44 020 7919 7873.

0167-4870/$ - see front matter
doi:10.1016/j.joep.2005.06.015

1. Introduction

As crawling gives way to toddling and then striding, a child may move more steadily through the physical world. So too, improvements in her ability to mentalise – that is, attribute, understand and manipulate mental states such as beliefs, feelings, thoughts, intentions and deceptions – allow her to navigate deeper into the cross-currents of the broader social world. For individuals with autistic spectrum disorders there exists a fundamental difficulty in mentalising and social life is a series of strong headwinds, uncertain tacks, and treacherous eddies. Specifically, individuals with autistic spectrum disorders (ASD) fail to understand not only that others have minds, but also that other minds have different thoughts, and that behaviour is determined by mental states. Thus, individuals with autism are considered to lack a 'theory-of-mind' (Baron-Cohen, Leslie, & Frith, 1985; Baron-Cohen, Tager-Flusberg, & Cohen, 1993, 2000).

While difficulties with theory-of-mind are widespread in individuals with autism, certain of these individuals appear able to acquire some degree of theory-of-mind understanding. In a meta-analysis, Happé (1995) reported that children with autism required a higher level of verbal ability, as measured by the British Picture Vocabulary Scale, in order to pass simple theory-of-mind tasks (the Sally–Anne and Smarties false belief tasks), than did normally developing 3- and 4-year olds or children with learning disabilities but without autism. It may be that the high verbal ability of these individuals allowed them to 'hack out' a solution to the tasks (Happé, 1995). Consequently, depending on a situation's degree of social complexity, the responses of such individuals with autism may be adequate and yet odd, or simply inapt and inept.

The limited ability of individuals with autism to compensate for their deficit in theory-of-mind has been demonstrated in two recent experiments. Castelli, Frith, Happé, and Frith (2002) asked participants to provide verbal interpretations of animations of two moving triangles. Each animation was scripted to show random, goal-directed or mentalising movements. Compared to their normally developing peers, individuals with ASD made fewer and less accurate interpretations only of the animations that evoked mentalising, for example when two triangles bounced up and down together in glee. A much more natural interaction was viewed by participants in Klin, Jones, Schultz, Volkmar, and Cohen (2002) study. While their eye movements were tracked, these participants watched dramatic scenes from a famous Hollywood movie. Individuals without autism focused mostly on the eyes of the actors; individuals with autism fixated mainly on the mouths. In contrast, when the scene showed a character reaching for a gun, the eyes of the individuals with autism moved directly to the object while those of the matched controls lingered on the actor's face, presumably to gather clues about his intentions. Thus, there is striking evidence that individuals with high-functioning autism read minds differently from their peers. While their performance on standard laboratory tasks of theory-of-mind (generally false belief) can be good, they are severely developmentally delayed in acquiring such ability, produce unusual explanations of their theory-of-mind understanding, perform poorly on advanced tests of theory-of-mind, and do not show the same spontaneous reactions to naturalistic task demands as their non-autistic peers.

All the studies above placed the participants in the role of an observer of another person or of a social interaction. Clearly, individuals with ASD experience difficulties with mentalising in this setting, but little has been documented empirically about the implication of this difficulty in their daily interactions with others. What happens when the person with autism is not spectator but participant, not outside but inside the interaction? The aim of the current study was to pursue this question by investigating the performance of both younger and older children, some of whom were normally developing and some of whom were diagnosed with autistic spectrum disorders, using a series of games.

Tic-tac-toe (also known as noughts and crosses), as any 10-year-old can testify, is a fairly trivial game, in part because the optimal strategy is so easily learned (capture the centre square with your "x" or "o") and is largely impervious to the behaviour or characteristics of one's opponent. However, when a game becomes more complex and its strategy space grows, optimal strategies are harder to calculate, errors are more readily made, and signals are more easily sent. In these games, the ability to mentalise may help a player anticipate where the other's memory, information, and calculation are limited and what strategy she is likely to employ. A well-drawn mental model may also distinguish when the counterpart's surprising move is a mistake or a bluff or a trap.

The participants in the complex games of poker and combat testify to the value of mentalising. A high-stakes card player, whose winnings depend on his abilities to deceive his competitors and to dispel others' bluffs, claims (with a bit of grandiosity), "A man's character is stripped bare at the poker table. If the other players read him better than he does, he has only himself to blame. Unless he is both able and prepared to see himself as others do, flaws and all, he will be a loser in cards, as in life" (Holden, 1990). Similarly, leaders on the battlefield rely on mental models of their counterparts to construct strategies and decipher tactical movements (see, for example, the memoirs of Grant, 1999/1885; Rommel, 1953).

While these zero-sum games are significant, it is principally through mixed-motive games that researchers have analyzed social and strategic interactions among humans, among other primates, birds and other vertebrates, insects, plants, genes, corporations, countries, political parties, classes, voters, etc. In the laboratory, three mixed-motive games in particular, the Prisoner's Dilemma, the Ultimatum game, and the Dictator game have been employed to study varying levels of cooperation and competition, concern for fairness, self-interest and altruism among experimental participants. Thousands of studies have been conducted of these three critical games, and participants have usually manifested a cooperativeness greater than that predicted by the models of strict, rational self-interest historically prominent in both biology – natural selection at the individual level – and economics – homo economicus (for reviews, see Camerer, 2003; Sally, 1995).

It is not necessary to mentalise in order to play a mixed-motive game or cooperate within such a game. For example, viruses competing to infect and reproduce in the same set of host cells are "playing" (insensibly and non-mentally) a Prisoner's Dilemma game in which an inability to cooperate lowers the fitness of each phage (Turner & Chao, 1999). Furthermore, biologists have discovered relatively sophisticated and accommodating strategies adopted by animals without the cognitive ability to form mental models – rotating responsibilities for approaching possible predators among stickleback fish (Milinski, 1987), reciprocal food sharing by vampire bats (Wilkinson, 1984), and the coordinated capture of grasshoppers and moths by certain species of spiders (Pasquet & Krafft, 1992). Evolutionary game theorists assume that the tactics of these creatures are encoded in their genes, that conspecifics are paired with a certain frequency in strategic interactions, and that fitness and offspring accrue to the genes and the individuals with the more robust strategy. So, vampire bats share food reciprocally because repeated interactions among colony-mates make that strategy more "profitable" than one based on always taking food and never giving. These bats are not analysing, rationalising, mentalising, or improvising: they are simply following an established evolutionary rule in a familiar situation.

Nevertheless, among humans mixed-motive games do seem to require that players signal and interpret intentions and develop some theory of the other's mind (Schelling, 1960). In the parlance of economics, then, mentalising and rule-following are substitutes in the same way that butter and margarine, aluminium and tin, work and the national lottery, cars and public transportation, and wisdom and information are. Few substitutes are perfect for all applications. A chef may be indifferent about which spread is applied to her morning toast, but would refuse to use oleo in even the simplest pastry, as butter makes a distinctly better crust. Similarly, the key research question we raise is, do mixed-motive games require specific mentalising abilities or will some form of rule-following suffice? Does the development of theory-of-mind promote cooperation, generosity, and fairness, or does it foster treachery and selfishness?

2. Experiment one – Prisoner's Dilemma

Without question, the Prisoner's Dilemma (PD) is the most thoroughly studied game in the social and biological sciences because it captures the essence of a frequent social quandary, namely, that what is good for the group may differ from what is good for each individual. Despite the gloomy Nash equilibrium of mutual defection and the predicted triumph of narrow self-interest over altruism, humans and other creatures often "solve" the PD and achieve mutual cooperation. Theorists have distinguished two principal means which aid mutual cooperation: kinship and reciprocity. It makes evolutionary sense for close kin to sacrifice on each other's behalf in order that their shared genes may prosper, and so brothers and sisters, rather than third cousins, will cooperate in the PD (Hamilton, 1964). Kin recognition depends on cognitive mechanisms able to read and react to perceptible clues of relatedness (Krebs, 1987), and vestiges of these prehistoric mechanisms translate into our modern predilection for people who are proximate, similar and familiar (Sally, 2000). Hence, in laboratory experiments, cooperation is more frequently seen when participants are near each other, can see each other, have information that they share tastes and opinions, like each other, and have interacted previously (Sally, 1995, 2000).

Finally, if the PD is repeated, a population of reciprocal altruists whose strategy is based on doing this round what the opponent did the previous round, can multiply and survive by reacting harshly to full-time defectors and nicely to full-time co-operators and to each other. Axelrod (1984) conducted a famous computerized tournament of repeated PDs in which the winner was this so-called tit-for-tat strategy.

Tit-for-tat (TFT) is a simple rule that requires absolutely no mentalising. One unilaterally cooperates in the initial round and then observes what move the counterpart actually makes so that it can be reciprocated in the next round. Intentions play no role here. Moreover, this type of player plays TFT with his brother and with a stranger, with a fellow club member and with an enemy, with a counterpart who promises to cooperate and with one who vows to defect. Such a rule-follower is relatively unaffected by changes in the social context of the game and is more likely to miss other changes in a game that occur below the surface. By contrast, kinship-based altruism is more closely linked to mentalising, both because the sharing of mental states is a sign of similarity and relation, and because the presence of those who are close, similar, and familiar is more likely to trigger an individual's theory-of-mind (Sally, 2001). Hence, the identity of the opponent makes a significant difference in the likelihood that a player chooses to cooperate. If the opponent's identity is changed from human to machine, even though the pre-programmed strategy is unaltered, experimental participants are far more likely to cooperate with the former rather than the latter: 58% versus 41% in Abric and Kahan (1972), and 59% versus 31% in Kiesler, Sproull, and Waters (1996). These two results suggest that the more "human" a counterpart is to a player, the more likely that player is to cooperate in the PD.

The degree of "humanness" seen in the opponent is a function not only of the opponent's identity but also of the identity and cognitive abilities of the perceiver. Chief among these abilities is mentalising. Insofar as theory-of-mind develops throughout childhood (Astington, 1994), one would expect, then, that older children would cooperate more than younger ones. Indeed, Fan (2000) found that 9- and 11-year-old children cooperated more frequently in a 10 round repeated PD than did 7-year olds. This study was less concerned with accounting for the choices of children than with the effects of moral suasion and instruction on the promotion of cooperation in the PD (see also Matsumoto, Haan, Yabrove, Theodorou, & Cooke Carney, 1986). Another study of young children (aged 3–10 years) and the PD showed that older children were more able and willing to pay attention to their opponent's interests than younger children, although this was only true if doing so would help improve their own outcome (Perner, 1979). Other developmental findings have shown that the frequency of prosocial behaviour increases throughout childhood (Eisenberg & Fabes, 1998). In the PD, cooperation, as opposed to defection, is clearly the prosocial move.

In our experiments the performance of normally developing children (aged 6, 8, and 10 years) as well as high-functioning children with autistic spectrum disorders was investigated on three versions of a PD game. In the PD games the nature of the opponent against whom a participant played was manipulated – each participant played the game against a human and a computer opponent. Furthermore a third manipulation of the PD was included in which participants played against a human opponent where the instructions of the game encouraged participants to cooperate, rather than compete with their opponent (encouraged cooperation PD).

Accordingly, based on the current understanding of theory-of-mind within developmental psychology and the theories and empirical results most frequently cited within behavioural economics, a number of hypotheses about the decisions of our participant groups were generated. With respect to the normally developing sample a greater degree of cooperation in the PD among the older participants was expected. Given the difficulties of individuals with autistic spectrum disorders in the area of mentalising, and the suggestion that mentalising is involved in a participant's choice on the PD, it was predicted that children with ASD would cooperate less when playing against the human opponent than their normally developing counterparts would. Since individuals with autism have been shown to have difficulty only representing the "mind" of a human and not a machine (Leekam & Perner, 1991; Leslie & Thaiss, 1992), a smaller difference in the cooperation rates between the normally developing children and those with ASD when playing the PD against a computer opponent was expected. With respect to the encouraged cooperation game, the specific instructions negate the need to predict the intentions of the counterpart, and so, individuals with and without autism should manifest similar, elevated rates of cooperation.

2.1. Method

2.1.1. Participants

A total of 69 children took part in the study. Participants were tested individually by an experimenter and – for the experimental tasks – a confederate, either in their school or at the Institute of Cognitive Neuroscience, UK. All were native English speakers. Participant details are shown in Table 1. All participants with ASD had been diagnosed formally with either autism or Asperger syndrome prior to the study. In addition, a checklist was completed by the experimenter and the confederate for approximately one-third of the participants. This checklist was based on observation

Table 1
Participant details

	Normally developing participants			ASD participants
	6 years	8 years	10 years	
N	14	19	18	18
Male (female)	8 (6)	10 (9)	10 (8)	16 (2)
Calendar age (year·month)				
Mean	6.7	8.5	10.6	10.6
Std. dev.	.2	.3	.3	3.1
Range	6.04–6.11	8.00–8.11	10.03–10.11	6.0–15.0
VIQ				
Mean	109.22	103.22	109.3	96.29
Std. dev.	10.87	16.35	11.92	33.7
Range	98.4–133.3	75.1–136.5	89.8–129.6	63.2–211.9

and related to the key characteristics of the disorder and was used as an aid to confirm an individual's diagnosis and therefore the results will not be reported here.

Ethics approval for the study was granted by the National Autistic Society (UK) and by University College London (UK). Parental consent was required for child participation in the study and the informed consent of all participants was sought.

2.1.2. General ability

General ability levels were assessed using the British Picture Vocabulary Scale (BPVS-II, Dunn, Dunn, Whetton, & Burley, 1997). Mean ability levels fell within the normal range in all groups (see Table 1).

2.1.3. Theory-of-mind

First-order false belief understanding was assessed in all participants using the Sally–Anne task (Baron-Cohen et al., 1985; Wimmer & Perner, 1983), and second-order false belief understanding was assessed using the Birthday Puppy story (Sullivan, Zaitchik, & Tager-Flusberg, 1994). The former test involves a vignette in which two girls are in the same room. Sally, having finished playing with a ball, places it in a basket and leaves the room. Anne takes the ball and moves it from the basket to a box that happens to be in the room also. Upon Sally's return, the key question is, "Where does she look for her ball?" An error in theory-of-mind is revealed if the child answers, "the box", since Sally is unaware of Anne's covert actions. While this task involves first-order mentalising, i.e., what are the contents of Sally's mind, the Birthday Puppy story is more complicated and asks whether the mother of a little boy who has been snooping around in the basement behind her back, believes that he is unaware of his surprise birthday gift – a puppy. The results shown in Table 2 reproduce the general finding in the literature: younger children have an immature theory-of-mind, and children with ASD have reduced mentalising capabilities. Specifically, children with ASD were impaired in both first- and second-order false belief tests in comparison to their normally developing peers [first-order false belief, $\chi^2(1) = 18.62, p < 0.001$; second-order false belief, $\chi^2(1) = 9.87, p < 0.01$].

2.1.4. Game theory tasks

The experimental game theory tasks were presented on a laptop computer, the screen of which was easily visible to both the participant and the confederate who sat side-by-side (the participant to the right), and the keyboard of which was divided by a partition hiding half the keys from each person. Participant responses were recorded on-line for later analysis. The instructions for each task were presented on

Table 2
Percent of each group passing first- and second-order false belief tasks

	Normally developing participants			ASD participants
	6 years	8 years	10 years	
First order	100	100	100	66.67
Second order	71.43	94.12	100	55.56

the computer, although these were verbalised by the experimenter to ensure that participants understood each task. In all tasks, players were told that they must try to win as many points as possible and that the total points won on all games would be exchanged at the end of the testing session for stickers. The greater the number of points won, the greater the number of stickers at the end of the test session. This incentive structure reflects the experience that non-monetary rewards are more appropriate when testing children (Fan, 2000; Harbaugh, Krause, & Liday, 2003). Two different trial games were played, before the experimental tasks, to ensure that the principle of the experimental tasks, the response methods, and the incentive structure were familiar to participants.

Three versions of a Prisoner's Dilemma task were presented to the participants, and 16 rounds were completed in each version, although participants were not told explicitly that this would be the case. This method allowed comparison of participants' strategy choice on the first round of the game as well as over all 16 rounds of each game.

2.1.4.1. Human opponent. A circle and triangle were shown on the computer screen. Each player (participant, confederate) independently chose either shape having been told how the points would be awarded (see Table 3). This information was outlined to the players verbally and was also presented on the tabletop in front of the players throughout the game. After players had made their choice by pressing the appropriate key on the keyboard, the computer presented the choices made and the number of points won. The confederate followed a tit-for-tat (TFT) strategy, always making the cooperative choice – triangle – on the first round and then copying the participant's strategy on the previous round for rounds 2–16. The participant was not told that there would be more than one round of the game.

2.1.4.2. Computer opponent. This version of the task was designed to investigate whether there was a difference in spontaneous strategy used when playing against a human or computer opponent. The task was identical to that described above with the computer replacing the human opponent. The computer was programmed to respond using the same strategy as the human opponent (cooperation on the first round followed by TFT), although this information was not given to participants. When playing the human and computer opponent PD games, defection (i.e., competition) is the equilibrium choice. Empirically, a wealth of studies and broad reviews of the literature (Dawes, 1980; Sally, 1995) suggest an expected cooperation rate of

Table 3
Points awarded to the participant and confederate in the Prisoner's Dilemma games

Confederate	Participant	
	Cooperate (triangle)	Compete (circle)
Cooperate (triangle)	3, 3	1, 4
Compete (circle)	4, 1	2, 2

Note: The players did not see the terms cooperate and compete. The terms confederate and participant were replaced with the appropriate names of the two players in each case.

approximately 20% in the first round and then approximately 40% overall as the repeated nature of the tasks is implicitly recognized. In both games a greater degree of competition than cooperation would be expected, especially when the opponent is a computer.

2.1.4.3. Encouraged cooperation. This task was identical to the PD with human opponent except that the points won on each round of the game were combined for the two players and divided equally at the end of the game. The confederate continued to follow a TFT strategy, always making the cooperative choice of triangle on the first round of the game and then copying the participant's strategy on the previous round for rounds 2–16. When both players chose triangle, six points were won collectively; circle, four points were won collectively; and in cases where each player chose a different shape, five points were won collectively. Note that with this payoff structure, cooperation becomes the dominant move: irrespective of what the other chooses, a player gains more points by cooperating. However, this dominance is somewhat obscured by the visible payoff matrix and the confederate's continued retaliatory response to the participant's choice of circle, and hence, the game is akin to a bluff – it is a PD on the surface and a purely cooperative game in reality.

The order of play against the human and computer opponents was counterbalanced across the participants, with the encouraged cooperation task being completed last in all cases.

2.2. Results and discussion

The mean number of cooperative responses across the 16 rounds of each PD task is shown for the normally developing 6-, 8- and 10-year olds and the children with ASD in Fig. 1. Two facts are immediately apparent: first, our participants were significantly competitive. Based on the literature cited earlier, we would have expected an average of six to nine cooperative moves in this setting, and our participants were well under that number in both the human and computer opponent versions. Second, the differences among the various age groups and the normally developing children and those with ASD are minimal. This similarity is confirmed by a repeated measures ANOVA with one between factor (participant group) and one within factor (PD game type). For performance across the 16 rounds of each game, there was a significant effect of game type [$F(2,63) = 30.49$, $p < .001$], indicating reliably higher levels of cooperation in the encouraged cooperation version of the task in comparison to both the human- and computer-opponent versions. However, there was no significant difference in the level of cooperation seen across the groups [$F(3,63) = 0.38$, $p > .1$]. Hence, when considering the results by age group and by diagnosis of ASD, none of our initial hypotheses are confirmed: older children are not more cooperative, the normally developing children are not more competitive with the computer, the children with ASD are not more competitive with a human counterpart, and not all subjects cooperate fully in the third, payoff-altered version of the PD.

The same overall cooperation rate could disguise a number of very different patterns. For example, cooperation could start off very high initially and then decay, or

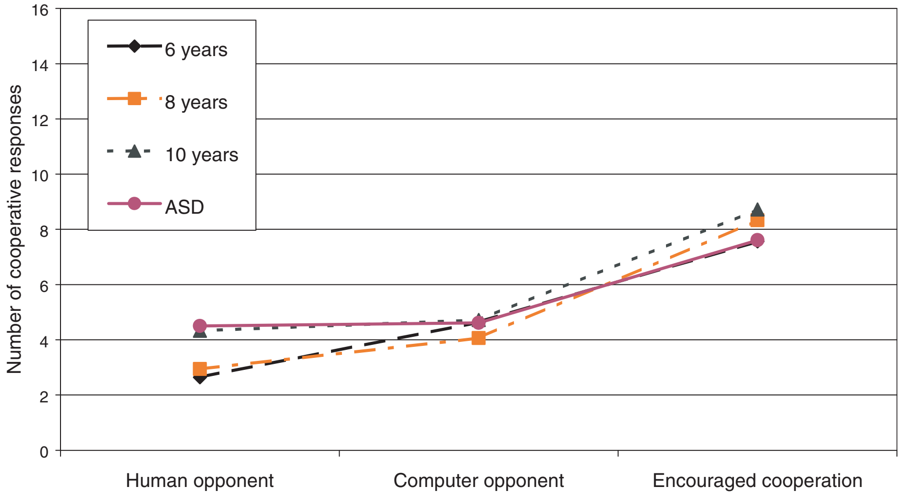

Fig. 1. Average cooperation rate over 16 rounds by group.

distrust could be prevalent at first with mutual cooperation emerging as TFT strategies are identified and synchronized. Moreover, given that the participants were never told explicitly about the number of rounds, cooperation in the first round was less likely to be motivated by reputation or reciprocity. Fig. 2 reveals that the 6-year olds generally were less likely to cooperate initially. Accordingly, this youngest group increased their average cooperation rate over the subsequent 15 rounds, while the

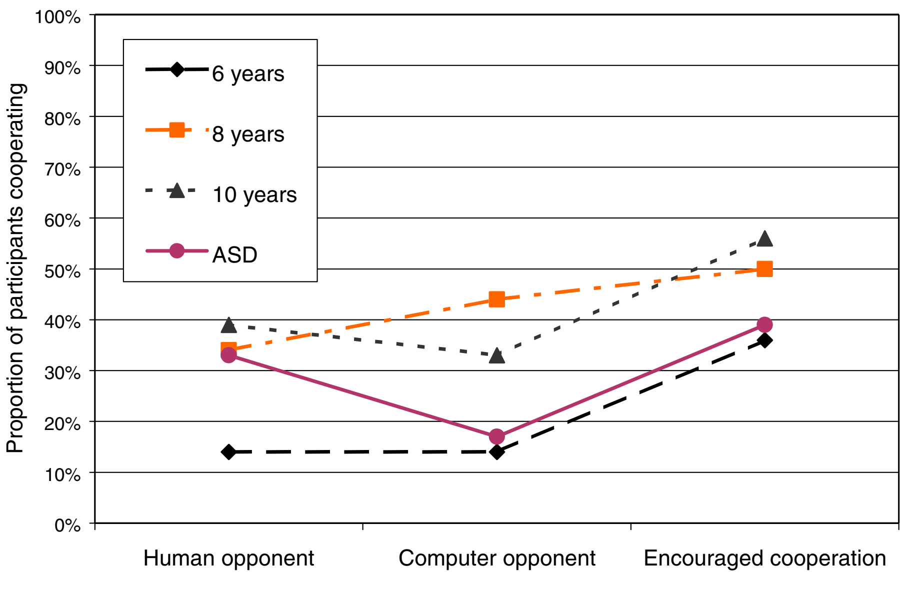

Fig. 2. Cooperation rate in initial round by group.

older normally developing children became more competitive. The children with ASD, in direct opposition to our hypothesis, were more competitive when they first confronted the computer opponent.

By comparing the choice made on the last round of the second PD game played (either against the human or computer opponent) with the choice made on the first round of the encouraged cooperation version of the PD (rounds 32 and 33 respectively), the extent to which individuals recognised the switch in the rational choice to be made according to the new task demands was considered. The data were investigated for each group separately using a series of Wilcoxon tests to compare the level of cooperation on each pairing of the PD games. There was a marginally significant difference between the choice of shape made in rounds 32 and 33 when the 6-, 8- and 10-year olds were considered separately [6 years, $Z = -1.63$, $p < .1$; 8 years, $Z = -1.73$, $p < .05$; 10 years, $Z = -1.0$, $p > .1$], although when combined, the normally developing children showed a significant difference between the choice of shape in the two rounds [$Z = -2.5$, $p < .01$]. Overall, then, the normally developing children adjusted immediately to the changes in the rules and payoff structure of the game. By sharp contrast, for the children with ASD, there was absolutely no significant difference between the choice of shape in the two rounds under consideration [$Z = .0$, $p > .1$]. It is conceivable that this difficulty in switching strategies and responding to the new payoff matrix is related to a general tendency in ASD toward rule-oriented behaviour and perseveration (Perner & Lang, 2000).

Taken together these results show that although there was no significant difference between the average degree of cooperation elicited in the individuals with ASD and controls when all rounds of each game were taken together, more subtle differences were evident in the detail. These differences and similarities can be more fully explained with an explicit analysis of the role of mentalising in fostering or retarding cooperation in these games.

In order to establish whether there was a relationship between mentalising ability and choice of strategy when playing each version of the PD, a comparison was made between performance on the second-order false belief test and the degree of cooperation evidenced in each game for the child participant groups. A repeated measures ANOVA with two between factors (group; second-order false belief performance) and one within factor (PD game) was applied to the data for number of cooperative moves (children with ASD versus all normally developing children together). There was a significant effect of game [$F(1,62) = 19.72$, $p < .001$], as described previously. The effect of false belief approached significance [$F(1,62) = 3.72$, $p = .058$], suggesting that second-order false belief passers had a tendency to be more cooperative than second-order false belief failers, irrespective of whether or not an individual was diagnosed with ASD. A child who failed this false belief test (i.e., saying the mother knew her son knew about the surprise birthday puppy) presumably also had a tough time understanding what the counterpart's beliefs about the child's own intentions were in the PD. Generally, a player wants to be cooperative only if she forsakes narrow self-interest and if she can anticipate that the other will cooperate (Kiyonari, Tanida, & Yamagishi, 2000). The target of this latter anticipation is, of course, a similarly conditional, cooperative expectation, and hence, the importance of a theory-of-mind.

None of the normally developing children failed the first-order false belief task, while a third of the children with ASD did so. To determine the relationship between this fundamental mentalising capacity and cooperation, the performance of the children with ASD was analysed with a one between factor (first-order false belief performance) and one within factor (PD game) repeated measures ANOVA. There was a significant effect of the game [$F(1,16) = 2.13$, $p < .05$], as described previously. Passing the false belief task significantly decreased cooperation across the games, mean number of cooperative responses, 4.72 and 7.27 for false belief passers and failers, respectively [$F(1,16) = 4.46$, $p < .05$]. There was no interaction between game and task performance [$F(2,16) = 2.01$, $p > .1$].

This result is the opposite of the one reported for second-order false belief: an acutely malfunctioning theory-of-mind enhanced cooperation. The children who could not decipher the first-order false belief task (i.e., they respond (mistakenly) that Sally believes the ball is now in the box) cooperated, on average, in approximately half the trials of each of the PD game versions. One of the ways a player could generate a cooperation rate of 50% is to simply choose randomly on each round without any regard for the decisions of the counterpart. If this form of decision-making were employed, the player would be equally likely to choose the circle or the triangle in a round regardless of which shape the opponent chose on the previous round, in other words, random reciprocation. (The tit-for-tat strategy, by contrast, assures that circle follows the other's circle and triangle, triangle with certainty.) For each of the children with ASD, the conditional response rates to the opponent's cooperating or defecting in the previous round were calculated. The hypotheses that these conditional response rates were equal to 50% were tested for those children who had passed or failed the Sally–Anne task. For the ASD children with first-order false belief troubles the hypothesis of random reciprocation could not be rejected in three of four instances (computer opponent: cooperation after cooperation $t(5) = .0$, $p > .1$, defection after defection $t(5) = 1.42$, $p > .1$; human opponent: cooperation after cooperation $t(5) = .44$, $p > .1$, defection after defection $t(5) = 4.07$, $p < .01$). In this last case, defection was reciprocated a little more than 70% of the time. By contrast, those who passed the first-order false belief test did not respond randomly in any setting (computer opponent: cooperation after cooperation $t(12) = -3.19$, $p < .01$, defection after defection $t(12) = 4.83$, $p < .001$; human opponent: cooperation after cooperation $t(12) = -3.71$, $p < .005$, defection after defection $t(12) = 6.23$, $p < .001$). This sub-group of test passers was more likely to reciprocate defection and to exploit cooperation. The evidence, then, strongly suggests that a first-order theory-of-mind is necessary to respond in a strategic fashion in the PD.

2.2.1. Summary

The broad similarity across age groups and between normally developing children and those with autistic spectrum disorder suggests that having immature mentalising capabilities (as in the case of the 6-year olds) or a hindered theory-of-mind (as in the case of the children with ASD) was not a major impediment in this version of the PD. We know that there are many social games – conversation, joint attention, pretend play, friendship – in which the difference between normally developing children and those with ASD are glaring. Clearly, then, this abstract version of the PD did not embody the social complexity of other forms of human interaction.

Nevertheless, mentalising ability aided strategic behaviour irrespective of the presence or absence of autism. Those children with ASD with very ineffectual theory-of-mind skills responded largely randomly to the moves of their human and computer counterparts. Among all the children, the ability to pass the more complex theory-of-mind task either intuitively or through conscious cognitive effort was correlated with greater levels of cooperation. Such a correlation may be driven by the need to perceive the opponent's intention to cooperate or reciprocate, a perception obscure to those without sufficient mentalising skills. These results are the first direct demonstration of the effects of mentalising ability on strategic behaviour in games.

In the PD, the issues of cooperation, generosity, and retribution are interwoven with uncertainty about the counterpart's intentions and actions and with a multiplicity of rationales for a given action in each round. These interweavings may be clarified by a second study, in which an alternative form of strategic interaction – bargaining – was investigated.

3. Experiment two – bargaining

Nature, an innumerate bank teller, an overstocked shelf, gravity, a misplaced wallet, an academic researcher, or some other dea ex machina might endow an individual not only with a prize but also with a companion and with a decision about sharing. In the Dictator game, the choice is how much of the prize to *grant* to the other party who is bound to accept the grant. In the Ultimatum game, the choice is how much of the value to *offer* to the other party who then may accept the offer or refuse it. Acceptance achieves the suggested division; refusal results in the whole prize being withdrawn and both parties receiving nothing. For example, an experimenter might confer upon a participant 10 candies (Murnighan & Saxon, 1998), 10 marks or some other unit of currency (Güth & Tietz, 1990), or 10 points or tokens (Harbaugh et al., 2003). This endowed person would, as a dictator, decide how many of the 10 units, if any, to give to another person, and as the proposer in the Ultimatum game, how many to offer the other party knowing that a rebuff would wipe out the grant entirely.

The Ultimatum game represents, among other social situations, the possibility in a negotiation of one bargainer making a final offer to the counterpart and walking away from the table leaving the other to sign the deal or not. Experience and introspection tell us that in this setting such a dramatic proposal has a good chance of failing. However, in orthodox economics, such an ultimatum should work: the equilibrium, corresponding to the prediction based on rational self-interest, is that the responder should accept any offer greater than zero, and therefore, the proposer should offer the smallest possible positive amount. In fact, this equilibrium is rarely realized in any of the numerous assays that have been conducted in laboratories and field sites around the world during the last 20 years. The regular findings, rather, are the following: (i) the modal offer is between 40% and 50% of the whole prize, (ii) tiny offers are almost always rejected, and (iii) a majority of responders will refuse offers below a third of the total (see Camerer & Thaler, 1995; Güth & Tietz, 1990; Oosterbeek, Sloof, & van de Kuilen, 2004 for reviews).

The literature has focused on understanding why offers are so robust, and why responders are so rancorous. One possible explanation for the generosity of the proposers is that they have a certain taste for fairness and a preference for sharing some part, or even half, of their dowry. Forsythe, Horowitz, Savin, and Sefton (1994) asked participants to play an Ultimatum game and a Dictator game to test this hypothesis, since fairness considerations would dictate that proposers give the same amount in both games. However, these authors found that the proportion of equal split offers declined from 75% in a $10 Ultimatum game to 21% in a $10 Dictator game. The mean offer in a standard 10 unit Dictator game is between 20% and 25% (Rigdon, 2003), and in the Ultimatum game, 40% and 45%. Roughly, then, half of the typical proposer's generosity is driven by a taste for fairness and half by strategic considerations of the possible spite of the responder.

Subsequent Dictator experiments have shown that the taste for fairness can be heightened or slaked by the context of the game. The degree of selfishness among dictators is raised by allowing them complete anonymity, even from the experimenter (Hoffman, McCabe, & Smith, 1996) and by placing them in a business setting of buying and selling (Hoffman, McCabe, Shacat, & Smith, 1994); the frequency of altruistic grants is raised if the recipient is a charity (Eckel & Grossman, 1996) and if more personal characteristics of the recipient are identified (Bohnet & Frey, 1999; Charness & Gneezy, forthcoming). Subsequent Ultimatum experiments have shown that the strategic and fairness concerns of both parties move in predictable directions: when the proposed split of $10 is generated by a roulette wheel as opposed to another person, the mean minimal offer acceptable to responders is much lower – $1.20 rather than $2.91 (Blount, 1995); when a written note saying "I know you'd like more, but that's the way it goes" is attached to a small offer, a rebuff from the responders is more likely (Kravitz & Gunto, 1992); when the endowment is worth more to one side than the other (e.g., 50¢ versus 10¢ per chip) and the other side is ignorant of this fact, advantaged proposers are more likely to suggest an even split of the counting units rather than total value, and advantaged responders are more likely to reject fair splits of the underlying units to induce a more equitable split of surplus value (Croson, 1996; Kagel, Kim, & Moser, 1996).

The variants just described could be reflective of mentalising ability or social rule-following. For instance, a person might reject a small offer because she imagines that the proposer thinks she is unworthy, gullible, or dim-witted, or because she recognizes this game as a sharing situation which demands that greedy people be punished. Similarly, a munificent dictator might utilize her theory-of-mind to anticipate the disappointment of an unfunded recipient, or she might recognize the applicability of a sharing norm. A norm may substitute for mentalising: as the other driver in a narrow lane approaches, you need not read his eyes, thoughts, or intentions, you need only remember the locale and move to the left in the UK and to the right in the US. Mentalising may become necessary only if the interaction does not proceed as expected: when the other driver fails to move to the proper side, then you need to notice the direction of his gaze, the tenseness of his hands, and the expression on his face.

The evidence on the relative utilization of mentalising versus norm-following in the standard, anonymous Ultimatum game is mixed. On the one hand, Henrich et al. (2001) found that much of the variance in mean offers among 15 small-scale societies (ranging from 26% among the Machiguenga in Peru to 58% among the Lamelara in Indonesia) could be explained by two factors – the importance of cooperation in the society's economic production and the reliance on market exchange in daily life. These factors would seem to develop and mould norm formation rather than mentalising abilities. The authors themselves suggest that their participants applied the analogous (and varying) norms found in their societies:

> [W]hen faced with a novel situation (the experiment), they looked for analogues in their daily experience, asking "What familiar situation is this game like?" and then acted in a way appropriate for the analogous situation. For instance, the hyper-fair…offers (greater than 50 percent) and the frequent rejections of these offers among the Au and Gnau reflect the culture of gift-giving…, accepting gifts, even unsolicited ones, commits one to reciprocate at some future time to be determined by the giver (Henrich et al., 2001, p. 76).

A second experiment that both documented the under-utilization of mentalising and its potential impact is that of Hoffman, McCabe, and Smith (2000). Here, in a $10 Ultimatum game with a business context, an additional line was added to the instructions encouraging the proposer (i.e., seller) to strategize and read the mind of the opponent: "Before making your choice, consider what choice you expect the buyer to make. Also consider what you think the buyer expects you to choose". This encouragement caused the mean offer to rise to $4.17 from $3.71 in a control condition where no explicit mindreading prompt was given, and this increase suggests that proposers in the control condition were solving the game without fully employing their theory-of-mind.

On the other hand, studies in which an asymmetry in information between the offerer and responder is strategically exploited support the relative prominence of mentalising. Respondents in one experiment received a set sequence of 12 offers for either $1 or $2 out of a total surplus of $20 (Pillutla & Murnighan, 1996). There was a complicated information structure overlaid on the set of games: (i) during the first half of the sequence, none of the responders knew that the total prize was $20, making it difficult to deem an offer unfair; (ii) half of the responders knew that the (fictitious) offerers knew that the low offers were unfair and therefore could more easily attribute greedy intentions to them. (Note that this second manipulation depends upon a theory-if-mind both to understand the different contents of the other's mind and to respond emotionally.) After making each decision, participants reported how they reacted to the offer and how they felt. These verbal responses were coded for the degrees of unfairness and anger expressed. In relation to the current study there were two important findings: first, the manipulation relying on theory-of-mind was successful – participants who knew both that an offer was low and that the other knew it was low were far angrier. Second, this anger resulted in a greater frequency of rejection, and was more predictive of the likelihood of rejection than was the expressed degree of unfairness alone.

As any overtaxed parent can testify, children reject various ultimatums all day long and readily employ multiple notions of "fairness." There is a vast literature on prosocial behaviour among children, some of it emphasizing social rules and some, perspective taking. Numerous donation studies (similar to the Dictator game) have found that children are more generous when they have seen a model being generous (e.g., Harris, 1970; Wilson, Piazza, & Nagle, 1990). Within social learning theory (Bandura, 1986), a model affects the observer by directly representing the presence or application of a rule rather than by triggering a mediating internal process, suggesting that mentalising is less important in giving. On the other hand, a child's abilities to take the perspective of another person visually, emotionally, and cognitively are positively related to prosocial behaviour in most studies (Underwood & Moore, 1982), and in one specific study, two factors relying on a child's theory-of-mind, affective reasoning and sympathy, caused a large increase in donations to a needy person (Knight, Johnson, Carlo, & Eisenberg, 1994). Both sources of prosociality should become stronger as children grow up, and indeed, a meta-analysis by Eisenberg and Fabes (1998) found that sharing and donating became more prevalent from preschool through adolescence.

Two specific Ultimatum studies, however, showed a less clear developmental trend. Harbaugh et al. (2003) found that fourth and fifth graders made larger grants as dictators than did second graders or ninth graders. While these authors found that ultimatum proposals on average increased with grade level, Murnighan and Saxon (1998) found a non-monotonic pattern with kindergarteners offering more candies than third graders, who, in a game with a dollar at stake, offered fewer cents than did sixth graders, who were more generous than ninth graders. Finally, younger children in both studies were more likely to accept small offers. This mixture of results demonstrates that the offering and responding behaviour of children may be affected by the specific details of game presentation, and may reflect a general inconsistency in inference about social interaction. Kalish (2002) has shown that while children and adults will equally predict consistency in repeated physical events such as pumice floating in water, children will much more often predict that a person would behave differently in the future than he or she did in the past, for instance, preferring Bert tomorrow despite preferring Ernie today. If the reaction of the other party is inconsistent or unreliable, then it makes sense to accept whatever the current offer is, and to not be too strategic in formulating an offer.

By investigating the giving and receiving behaviour of children with and without ASD, the importance of mentalising for both generosity and consistency can be determined.

3.1. Method

The participants were the same as those included in the first experiment, with the addition of one 8-year-old.

3.1.1. Materials and methods

The testing set-up remained the same as that reported for the first experiment with participants being assessed in a quiet room, sitting at a laptop to the right of the confederate. Responses were recorded on-line for later analysis. Task instructions were presented on the computer and the experimenter verbalised them to ensure that participants understood each task. Players were told that they must try to win as many points as possible and that the total points won on the games would be exchanged at the end of the testing session for stickers. Two versions of a bargaining task – Dictator game and Ultimatum game – were completed by participants, with the Dictator game always played first.

3.1.1.1. Dictator game. The dictator (participant) was given 10 points and asked how much s/he wanted to give to the opponent, knowing that s/he would keep the remainder. Eleven cards were presented on the computer screen, outlining all possible permutations by which the points could be split, ranging from the dictator keeping all 10 points for her/himself to giving all 10 points to the opponent. The dictator made his or her choice and indicated this by selecting the relevant card on the computer screen. The choice that the dictator had made and the points allocated to both players were displayed. This process was repeated 16 times throughout the course of the game, with the participant acting as the dictator for rounds 1–4 and 9–12, and the confederate taking the part of the dictator for the remaining rounds. The participant was unaware that there would be more than one round of the Dictator game and that the confederate would also take a turn as the dictator. In the latter case the confederate allocated approximately the same amount of points to the participant as the participant had to her.

3.1.1.2. Ultimatum game. This game was essentially the same as the Dictator game but the counterpart had the choice to accept or reject the offer made to him or her by the proposer in each round of the game. The game started in the same manner as the Dictator game. Once the proposer had made an offer, the opponent indicated whether s/he accepted or rejected that offer. If the opponent accepted the proposer's offer, the points were divided as proposed (exactly as in the Dictator game). If the opponent rejected the proposer's offer, neither player received any points. The choices made by each player, as well as the points won, were displayed after each round of the game. The set-up of the Ultimatum game was identical to that of the Dictator game, with the participant acting as the proposer and the confederate as the opponent for rounds 1–4 and 9–12, and with the roles reversed for the remaining rounds. The participant was not told explicitly that there would be more than one round of the game and that the confederate would also take a turn as the proposer. As a responder, the confederate always rejected offers of less than an equal split (i.e., four or fewer points) and, as a proposer, she allocated approximately the same amount of points to the participant as the participant had to her.

3.2. Results and discussion

3.2.1. Offers made

The offers made by the participant to the confederate were expected to be lower in the Dictator game than in the Ultimatum game. The mean points offered to the confederate by the participant across the first four rounds of each game, and on the first

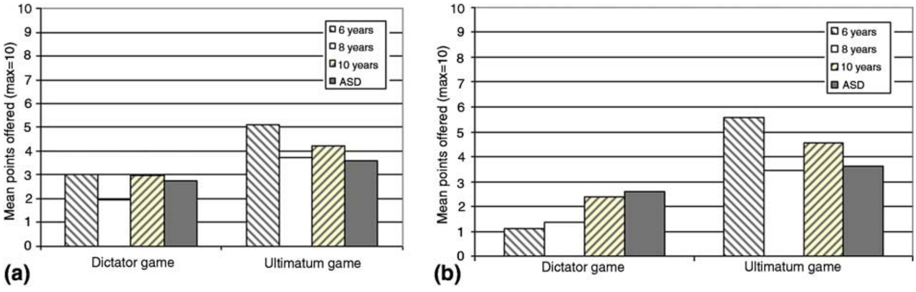

Fig. 3. (a) Average points offered to counterpart on first four rounds. (b) Average points offered to counterpart in first round.

round only, are shown for each group in Fig. 3a and b respectively. A repeated measures ANOVA with one between factor (group) and one within factor (game) was applied to the data for mean points offered. For performance across the first four rounds of each game, there was a significant effect of group [$F(3,65) = 3.70$, $p < .02$]. A series of Tukey tests revealed this difference to arise from higher offers being made by the six- in comparison to the 8-year-old children [$p < .05$]. There was a significant effect of game [$F(1,65) = 49.38$, $p < .001$], reflecting higher offers being made to the confederate in the Ultimatum game, and an insignificant interaction between group and game [$F(3,65) = 1.66$, $p > .10$]. These results correspond to the equivocal results of previous Ultimatum studies in that there was no clear trend in offer amounts across age groups. More work is needed to explain this finding within the prosocial development literature which shows that sharing and donating become more prevalent as children grow up (Eisenberg & Fabes, 1998). A tentative explanation for the differences may be that this task evoked exchange norms which are already fairly firmly implanted by the age of six.

We can begin to distinguish the groups of children when we look at what they do when they first encounter the bargaining games. A slightly different pattern of performance was seen in the analysis of the mean offers made to the confederate on the first round of each game [group, $F(3,65) = 1.03$, $p > .1$; game, $F(1,65) = 52.89$, $p < .001$; group by game, $F(3,65) = 4.18$, $p < .01$]. The interaction is driven largely by the 6-year olds whose initial proffered share increases by 44% as the game changes from Dictator to Ultimatum. In turn, the offers of the children with ASD, both at first and over the first four trials, grew only modestly as the receiver was enabled to say "no" in the Ultimatum game.

More evidence for the relative effects of autism on fairness and tactical giving appear when looking at the underlying distributions of offers. In particular, Fig. 4 displays the dispersal of first round ultimatum offers by all normally developing children versus those of the children with ASD. Visually, there appears to be a strong divergence. We used the powerful Epps and Singleton (1986) test to determine if these two samples of discrete data were likely to be from identical populations. This *CF* statistic, based on the empirical characteristic function, is asymptotically distributed as chi-square with four degrees of freedom and can be corrected for small samples. For the

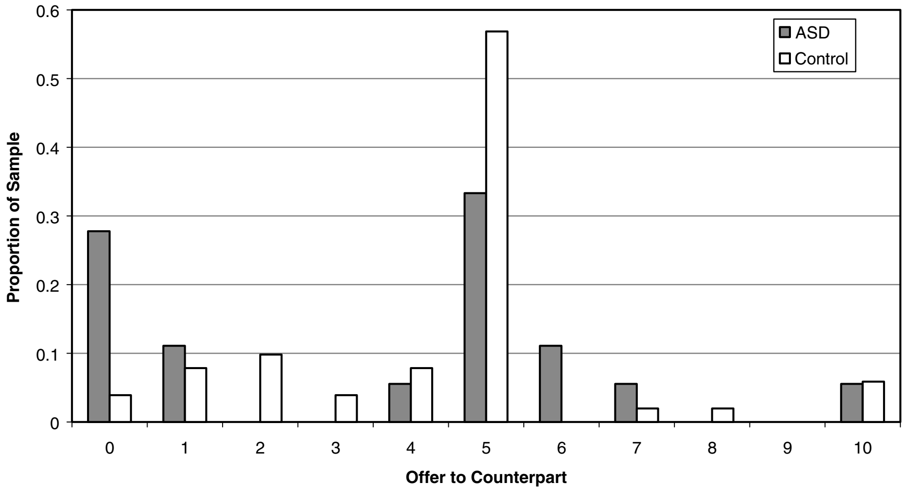

Fig. 4. Distribution of initial ultimatum offers by participants with and without ASD.

first round ultimatum offers, the null hypothesis of similar distributions is strongly rejected [$CF = 9.96$, $p < .05$]. Moreover, when comparing first round dictator offers there were no significant differences in the underlying distributions across the participants. Since the initial dictator grants were the same, it can be concluded that core generosity did not vary by age or with ASD. However, once an element of strategic anticipation was added by empowering the responder in the Ultimatum game, those diagnosed with autistic spectrum disorders diverged. A majority of these children seemed to employ one of two salient rules: cut the total in half, or keep it all, resulting in offers of five and zero, respectively. In contrast, more than half of the normally developing participants offered to share equally.

The direct effects of mentalising on first round offers among all the children can be seen in Fig. 5. The striking visual difference in the distribution of offers is strongly confirmed by an Epps-Singleton test ($CF = 9.64$, $p < .05$). Here, the majority of those children, both control and ASD, who failed the second-order false belief test offered the responder nothing or only one point, while the majority of those who passed the test offered an even split. (These distributions of initial offers have means whose distinctiveness approaches significance [$F(1,63) = 3.73$, $p = .058$].) A theory-of-mind that was effective due to intuition or synthetic construction was very helpful to the child making a reasonable offer in this game.

The extent of learning across rounds of the Ultimatum game was investigated, in light of the apparently different strategies of the individuals with and without ASD. The confederate consistently rejected offers of four points or fewer, so participants had the chance to learn, adjust and converge on the optimal offer of an even split. A comparison of the groups on a round by round basis was made on the Dictator and Ultimatum games for all eight rounds. A repeated measures ANOVA with one between factor (group) and two within factors (bargaining game, round) was applied to the data for the mean offer made by each group on all eight rounds of each game.

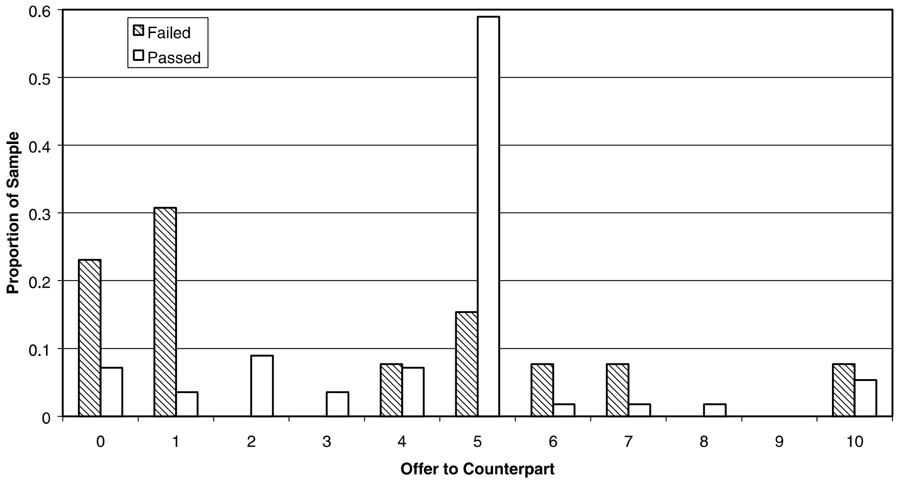

Fig. 5. Distribution of initial ultimatum offers based on second order false belief test result (all children).

There was a significant effect of game [$F(1,67) = 62.38$, $p < .001$], described previously and a significant interaction between group, bargaining game and round [$F(1,67) = 4.21$, $p < .01$]. This interaction highlighted divergence between the amounts offered to the confederate in the two games on round one in the normally developing children only. Most of the children seem to learn the game quickly: by the second round the offers of children with ASD and those of normally developing children are dispersed in statistically similar ways. This concurrence is confirmed by the insignificance of second-order theory-of-mind in an ANOVA of average offer over all eight rounds. The one group that remains distinct are the 6-year olds: even after 11 rounds of the game, in their final turns as proposers, their offers are scattered over all the possibilities and are distributed differently from those of the 8- and 10-year olds [$CF = 11.62$, $p < .02$] and those of the children with ASD [$CF = 8.29$, $p < .10$]. Again, this excess dispersion may be in keeping with the finding that younger children are less likely to predict consistency in the actions of other people, in this case, the responder (Kalish, 2002).

3.2.2. *Offers rejected*

One consistent finding of the bargaining studies discussed earlier has been that younger, normally developing children are more likely to accept smaller offers. Our data confirm this pattern: when faced with an offer of less than half of the 10 points, the 6- and 8-year old responders said yes 22.8% of the time, while the 10-year olds agreed only 10.6% of the time. Interestingly, this difference was caused not by varying reactions to tiny offers of one or two points, but by the younger children's willingness to accept four points (36% acceptance of these offers, compared to 14% for the 10-year olds). There was one additional bit of evidence on the relationship of age to offer rejection: three 6-year olds were the only participants to accept every offer made by the confederate.

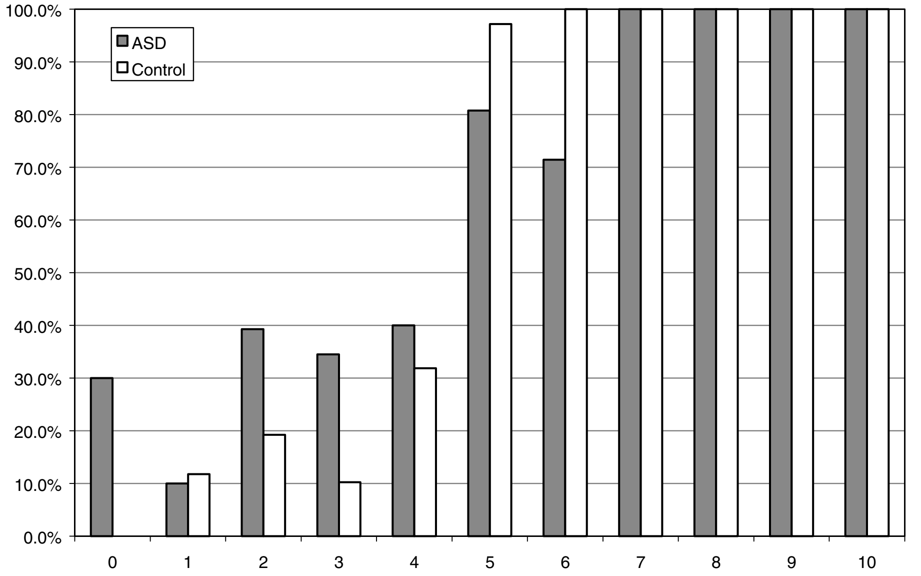

Fig. 6. Acceptance rates of different offer sizes.

A comparison of the respondent decisions of the control participants and those with ASD is shown in Fig. 6. Children diagnosed with ASD were much more likely to accept not only minority offers but piddling ones as well. Whereas the normally developing children refused 89% of the confederate's proposals of three or fewer points, the ASD children rebuffed only 68% of this level of offer. On the other hand, when the offer was fair or more than fair (five or six points out of 10), then the normally developing children almost never refused it (2% of offers), while ASD children were more likely to decline (21% of offers). Since the incentive structure was clear to all participants, the variation between the groups' respondent decisions is likely to be a manifestation of their relative abilities in discerning the intentions of the offerer. For the children with ASD these intentions are likely to be quite obscure and indecipherable.

3.2.3. Summary

The behaviour of the youngest normally developing children (aged 6 years) diverged most radically as the context of the game changed. Not only were they stingy in their first trial as a dictator, but they were, initially, overly beneficent as the ultimatum offerer, and they were unaffected by the repetition and learning of the game. For the older groups, repeating the Ultimatum game caused differences to dissipate and distributions of offers to converge. Children with ASD were distinguished from their peers in the initial round of the Ultimatum game but not the Dictator game. Children with ASD showed an equal predilection for proposals of nil or five in contrast to normally developing children who were most likely to just divide the prize in half. An ineffective theory-of-mind was apt to result in an initial offer of no more than one point, while deciphering the second-order false belief story correctly tended to lead to an equitable offer. The development of theory-of-mind skills may help the child first to recognize and act upon relevant norms of behaviour such as fairness, and later, to stretch those norms and improvise away from them when the situation calls for it.

A single experience as proposer was sufficient to counteract any deficits in mentalising in the succeeding rounds of the this Ultimatum game. Without question, the static nature of the payoffs and counterpart in this bargaining game is essential to the observed decay in the effects of mentalising. A more dynamic and realistic negotiation would probably continue to demand the application of an active theory-of-mind.

Finally, the rejection of offers did vary with age and with diagnosis: both younger normally developing children and those with ASD were more accepting of small offers. It is possible that the ability to perceive another's unfair intent or contempt is necessary to reject offers of 40% percent or less, and oppositely, the ability to sense regard is necessary to accept generous proposal. Hence, even in this very basic negotiation setting, the generation of an initial appropriate offer, the denial of an inapt or inequitable offer, and the acceptance of a fair proposal may depend on the bargainer's theory-of-mind.

4. Conclusion

We have made two contributions to behavioural economics and to the understanding of economic socialisation: first, we have introduced a new pool of experimental participants to the literature – those diagnosed with autistic spectrum disorder. One of the core deficits of autism involves the theory-of-mind, the ability to perceive and interpret the contents of another person's mind. Theory-of-mind is critical to many, if not most, forms of social interaction, and by serving as a partial contrast, those with ASD can illuminate what roles mentalising plays in the behaviour of non-autistic people. Second, by utilizing the false belief tests central to the study of autism in children, we have been able to discern some of the direct effects of mentalising ability on cooperation, fairness, generosity and strategic behavior. A minimally functioning theory-of-mind is necessary to make an intentional, non-stochastic choice in the repeated PD. A child's ability to pass the second-order false belief test, arising from either a functional theory-of-mind or effortful rule-following, was found to be positively related to the likelihood of cooperation in the three PD games and to perfectly fair ultimatum offers rather than very small proposals. Furthermore, participants with ASD accepted small offers more frequently.

While we believe these findings are significant, a larger question looms – why were there so many concurrences between the children with ASD and those who were normally developing? Such similarity does not mirror the behavioural differences observed in such individuals in similar real-life situations. One possible answer lies in the specifics of our protocol: the payoffs in the PD matrix may have been too subjectively close to motivate cooperation; the presence of the confederate, though necessary to help guide the participants with ASD through the lengthy testing session, may have been awkward for all subjects; the repetition of the bargaining games may have been more tiring than fruitful; within subject effects across games may not be completely controlled for; etc. These flaws should certainly be remedied in subsequent experiments.

A second answer is that a fully developed theory-of-mind may be sufficient but not necessary for basic strategic rationality, sometimes knowledge of basic rules of behaviour will suffice. Mentalising capability increased levels of cooperation in some settings and degrees of strategic behaviour in others. In other ways, however, we did not see the drastic discontinuities between the normally developing children and the children with ASD that we would see if the game involved reciting a joke, telling a story, creating an imaginary scene, appreciating an emotion, or gazing into another's eyes. Perhaps it is the on-line aspects of mentalising and mental flexibility that cause the greatest difficulty for high-functioning individuals with ASD in dilemma and bargaining situations in the real-world. These "games" are more suffused in social context with more distractions and fewer direct cues to guide behavioural choices than are those in the sterile laboratory where tasks can be seen to be abstract. If this answer is correct, then it reinforces the position that there is a significant discontinuity between games in the laboratory and games in the market, the office and the home, for those with ASD and those without. We are currently investigating this possibility.

Acknowledgements

This research was supported by funding from the UK's Medical Research Council (grant number G9716841). We gratefully acknowledge the willing participation of all individuals in this study. We are indebted to Sarah Griffiths, Zoë Fortune and Sakina Adam-Saib for substantial help with data collection and especially to Professor Uta Frith for invaluable support for, and discussion of the project.

References

Abric, J. C., & Kahan, J. P. (1972). The effects of representations and behavior in experimental games. *European Journal of Social Psychology, 2*, 129–144.

Astington, J. W. (1994). *The child's discovery of the mind.* London: Fontana.

Axelrod, R. (1984). *The evolution of cooperation.* New York: Basic Books.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

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: Oxford University Press.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (2000). *Understanding other minds. Perspectives from developmental cognitive neuroscience.* Oxford: Oxford University Press.

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attribution on preferences. *Organizational Behavior and Human Decision Processes, 63*, 131–144.

Bohnet, I., & Frey, B. S. (1999). Social distance and other-regarding behavior in Dictator Games: Comment. *American Economic Review, 89*, 335–339.
Camerer, C. F. (2003). *Behavioral game theory*. Princeton: Princeton University Press.
Camerer, C., & Thaler, R. H. (1995). Anomalies: Ultimatums, dictators and manners. *Journal of Economic Perspectives, 9*, 209–219.
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., & Gneezy, U. (forthcoming). What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior and Organization.*
Croson, R. T. A. (1996). Information in ultimatum games: An experimental study. *Journal of Economic Behavior and Organization, 30*, 197–212.
Dawes, R. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169–193.
Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British picture vocabulary scale* (2nd ed.). Windsor: NFER Nelson.
Eckel, C. C., & Grossman, P. J. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior, 16*, 181–191.
Eisenberg, N., & Fabes, R. A. (1998). Prosocial development. In W. Damon (Ed.), *Handbook of child psychology* (5th ed., Vol. 3, pp. 701–778). New York, NY: Wiley & Sons.
Epps, T. W., & Singleton, K. J. (1986). An omnibus test for the two-sample problem using the empirical characteristic function. *Journal of Statistical Computation and Simulation, 26*, 177–203.
Fan, C.-P. (2000). Teaching children cooperation – an application of experimental game theory. *Journal of Economic Behavior and Organization, 41*, 191–209.
Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior, 6*, 347–369.
Grant, U. S. (1999/1885). *Personal memoirs*. New York: The Modern Library.
Güth, W., & Tietz, R. (1990). Ultimatum bargaining behavior. *Journal of Economic Psychology, 11*, 417–449.
Hamilton, W. D. (1964). The genetical evolution of social behavior: I and II. *Journal of Theoretical Biology, 7*, 1–52.
Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*, 843–855.
Harbaugh, W. T., Krause, K., & Liday, S. (2003). *Bargaining by children*. University of Oregon Economics Working Paper No. 2002-4.
Harris, M. B. (1970). Reciprocity and generosity: Some determinants of children. *Child Development, 41*, 313–328.
Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review, 91*, 73–78.
Hoffman, E., McCabe, K., Shacat, K., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior, 7*, 346–380.
Hoffman, E., McCabe, K., & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review, 86*, 653–660.
Hoffman, E., McCabe, K., & Smith, V. L. (2000). The impact of exchange context on the activation of equity in ultimatum games. *Experimental Economics, 3*, 5–9.
Holden, A. (1990). *Big deal: A year as a professional poker player*. New York: Viking.
Kagel, J. H., Kim, C., & Moser, D. (1996). Fairness in ultimatum games with asymmetric information and asymmetric payoffs. *Games and Economic Behavior, 13*, 100–110.
Kalish, C. W. (2002). Children's prediction of consistency in people's actions. *Cognition, 84*, 237–265.
Kiesler, S., Sproull, L., & Waters, K. (1996). A prisoner's dilemma experiment on cooperation with people and human-like computers. *Journal of Personality and Social Psychology, 70*, 47–65.
Kiyonari, T., Tanida, S., & Yamagishi, T. (2000). Social exchange and reciprocity: Confusion or a heuristic. *Evolution and Human Behavior, 21*, 411–427.
Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Defining and quantifying the social phenotype in autism. *American Journal of Psychiatry, 159*, 895–908.

Knight, G. P., Johnson, L. G., Carlo, G., & Eisenberg, N. (1994). A multiplicative model of the dispositional antecedents of a prosocial behavior: Predicting more of the people more of the time. *Journal of Personality and Social Psychology, 66*, 178–183.

Kravitz, D., & Gunto, S. (1992). Decisions and recipients in ultimatum bargaining games. *Journal of Socio-Economics, 21*, 65–84.

Krebs, D. (1987). The challenge of altruism in biology and psychology. In C. Crawford, M. Smith, & D. Krebs (Eds.), *Sociobiology and psychology: Ideas, issues, and applications* (pp. 81–118). Hillsdale, NJ: Lawrence Erlbaum.

Leekam, S., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition, 40*, 203–218.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Evidence from autism. *Cognition, 43*, 225–251.

Matsumoto, D., Haan, M., Yabrove, G., Theodorou, P., & Cooke Carney, C. (1986). Preschoolers' moral actions and emotions in prisoner's dilemma. *Developmental Psychology, 22*, 663–670.

Milinski, M. (1987). Tit for tat in sticklebacks and the evolution of cooperation. *Nature, 325*, 434–435.

Murnighan, J. K., & Saxon, M. S. (1998). Ultimatum bargaining by children and adults. *Journal of Economic Psychology, 19*, 415–445.

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

Pasquet, A., & Krafft, B. (1992). Cooperation and prey capture efficiency in a social spider, anelosimus-eximius. *Ethology, 90*, 121–133.

Perner, J. (1979). Young children's preoccupation with their own payoffs in strategic analysis of 2×2 games. *Developmental Psychology, 15*, 204–213.

Perner, J., & Lang, B. (2000). Theory of mind and executive function: Is there a developmental relationship? In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds. Perspectives from developmental cognitive neuroscience* (2nd ed., pp. 150–181). Oxford, England: Oxford University Press.

Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes, 68*, 208–224.

Rigdon, M. L. (2003). Dictator games. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (pp. 379–382). London, England: MacMillan.

Rommel, E. (1953). In B. H. Liddell Hart & P. Findlay (Eds.), *The Rommel papers*. New York: Harcourt, Brace, Jovanovich.

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

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

Sally, D. (2001). Into the looking glass: Discerning the social mind through the mindblind. *Advances in Group Processes, 18*, 99–128.

Schelling, T. C. (1960). *The strategy of conflict*. Cambridge: Harvard University Press.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30*, 395–402.

Turner, P. E., & Chao, L. (1999). Prisoner's dilemma in an RNA virus. *Nature, 398*, 441–443.

Underwood, B., & Moore, B. (1982). Perspective-taking and altruism. *Psychological Bulletin, 91*, 143–173.

Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature, 308*, 181–184.

Wilson, C. C., Piazza, C. C., & Nagle, R. J. (1990). Investigation of the effect of consistent and inconsistent behavioral example upon children's donation behaviors. *Journal of Genetic Psychology, 151*, 361–376.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and the constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.