Insensitivity to social reputation in autism

Keise Izuma[a,1], Kenji Matsumoto[b], Colin F. Camerer[a], and Ralph Adolphs[a]

[a]Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125; and [b]Brain Science Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan

Edited by Lee D. Ross, Stanford University, Stanford, CA, and approved September 12, 2011 (received for review May 2, 2011)

People act more prosocially when they know they are watched by others, an everyday observation borne out by studies from behavioral economics, social psychology, and cognitive neuroscience. This effect is thought to be mediated by the incentive to improve one's social reputation, a specific and possibly uniquely human motivation that depends on our ability to represent what other people think of us. Here we tested the hypothesis that social reputation effects are selectively impaired in autism, a developmental disorder characterized in part by impairments in reciprocal social interactions but whose underlying cognitive causes remain elusive. When asked to make real charitable donations in the presence or absence of an observer, matched healthy controls donated significantly more in the observer's presence than absence, replicating prior work. By contrast, people with high-functioning autism were not influenced by the presence of an observer at all in this task. However, both groups performed significantly better on a continuous performance task in the presence of an observer, suggesting intact general social facilitation in autism. The results argue that people with autism lack the ability to take into consideration what others think of them and provide further support for specialized neural systems mediating the effects of social reputation.

Asperger syndrome | audience effect | dictator game

Concern for our own reputation affects how we behave in social situations. Our actions are strongly influenced by our belief that they may be seen and evaluated by others. Not only do people care about their reputation, but they also often try to manipulate what other people think of them through self-presentation or impression management (1), topics with a long history in social psychology. It is well known that subjects tend to behave in a more egoistic manner under guaranteed anonymity, whereas less anonymous situations increase prosocial behaviors (e.g., giving some benefit to others, adhering to a public standard or to social norms) (2). Prosocial behaviors can be elicited not only by the presence of real observers (3–5) but also by surprisingly subtle cues associated with being watched by others (6, 7), pointing to the powerfully automatic nature of reputation-based processing. Although altruistic behaviors toward nonkin may be present even in nonhuman primates (8), it is likely that only human altruistic behavior is affected by the presence of an independent third party (9). Representing our reputation involves thinking about what others think of us and thus requires some level of metarepresentation (10). The link between prosocial behaviors and reputation is also emphasized by theoretical and empirical considerations of how altruism might have evolved through indirect reciprocity (11, 12). Thus, reputation may be a unique and important aspect of our species that incentivizes self-interested individuals to conform to social norms.

Autism spectrum disorders (ASD) are a class of neurodevelopmental disorders with an estimated incidence near 1%, characterized in part by profound impairments in reciprocal social interactions (13, 14). The social difficulties encountered by people with ASD are especially striking in high-functioning individuals and are often the single most disabling component in everyday life. Although the core cognitive processes whose impairment might account for these behavioral deficits remain unknown, a large body of research has demonstrated impairments in social perception and social cognition in ASD (15), as well as abnormal structure and function in the associated neural structures (16). One leading hypothesis is that people with ASD are impaired in theory-of-mind abilities (the ability to represent what other people believe, even when those beliefs are false) (17, 18). However, deficits in theory-of-mind abilities are often difficult to demonstrate experimentally in high-functioning adults. Many prior studies used tasks in which subjects only observed social interactions but did not themselves engage in social interactions (17). Recently, several studies investigated social cognition in ASD while subjects were actually playing economic games with a human player (19–21). However, these studies found that the behavior of people with ASD during prisoner's dilemma games (20, 21) and trust games (19) was largely normal, and abnormalities only became apparent when analyzing neuroimaging data collected during the tasks (19). One recent study (22) found behavioral evidence that ASD may feature impairments in a metacognitive ability to recursively represent other's beliefs about one's own beliefs; however, in that study participants interacted with a complex computer strategy rather than with another person, leaving it unclear to what extent a specific deficit in social cognition might account for the findings. Taken together, the results to date suggest that high-functioning people with ASD may be impaired selectively in difficult or metacognitive aspects of social cognition but leave the precise nature of such a process unclear.

We hypothesized that individuals with ASD might have a specific deficit in taking into account their reputation in the eyes of others. This should be apparent in situations in which reputation works as an incentive for social behavior. The importance of testing reputation-based effects on social behavior in ASD is twofold: it may shed light on the specific processes that account for real-life social impairments in ASD; and it may provide a neuropsychiatric dissociation demonstrating the modularity of reputation-based processing in social behavior. It has been previously suggested that reputation-based effects could provide a sensitive measure of impaired social cognition in ASD (23) based on the findings of a neuroimaging study with high-functioning ASD (19).

In the present experiment, subjects performed a dictator game in which they were given an endowment of money and then presented with a series of choices asking them whether they wanted to donate variable portions of their money (or none) to a charity (Donation task; Fig. 1 *A* and *B*). Each subject participated in two conditions (in counterbalanced order): once alone in the room and once with an unfamiliar person sitting behind them and observing their choices (with a cover story to explain why the person was there; *Materials and Methods*). To control for nonspecific effects of the presence of another person, such as arousal or social facilitation (24), all subjects also performed a simple continuous performance task (Fig. 1*C*) in the presence/absence of the observer.

Author contributions: K.I., K.M., C.F.C., and R.A. designed research; K.I. performed research; K.I. analyzed data; and K.I. and R.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: izuma@caltech.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1107038108/-/DCSupplemental.

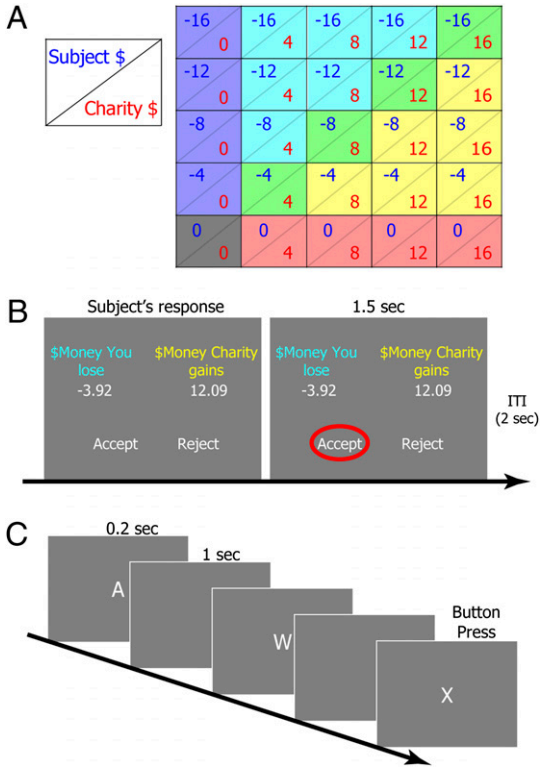

Fig. 1. Experimental tasks. (*A*) Design of Donation task. In each cell the number at top left indicates the amount of money subjects lose, and the number at bottom right indicates the amount of money the charity gains. Each design cell was implemented twice in each Presence and Absence session (i.e., 50 trials). (*B*) Example of a trial in the Donation task. A small random jitter was added to each monetary amount. After subjects decided whether to accept or reject the presented monetary transfer, their choices were highlighted by a red circle for 1.5 s. (*C*) Sequence of trials in the CPT. In each trial, a letter of the alphabet was presented for 0.2 s, and subjects were asked to press a button only when a letter "X" was presented. Subjects performed the task continuously for approximately 13 min.

Results

We compared a group of 10 high-functioning adults with ASD with 11 healthy controls matched on age, sex, and intelligence quotient (IQ) (Table S1). As seen in Fig. 2, when no observer was present, both control and ASD groups based their donations on how much money they would lose and how much money the charity would gain, showing normal preferences for their own monetary outcomes as well as normal social preferences for benefitting a charity. We quantified these findings with logistic regressions and found that decisions to accept donations depended significantly both on the cost to the subject and gain to the charity. For both groups (ASD and controls), monetary loss for subjects and gain for the charity had significant effects on choices (all $P < 0.001$). Both groups chose to donate less frequently as the amount of money they had to lose increased but chose to donate more frequently as the amount of money benefitting the charity increased (Fig. 2).

To test for the effect of an observer on donation decisions, a 2 (group; ASD or control) × 2 (observer; Presence or Absence) mixed ANOVA was first performed on the number of accepted donations during the Donation task. It revealed, as predicted, a significant interaction [$F(1,19) = 7.03$, $P = 0.02$] (Fig. 3*A*). Neither main effects of group nor observer were significant [all $P > 0.17$, nonsignificant (n.s.)]. Direct comparisons between Presence and Absence conditions within each group showed that whereas our observer manipulation was successful in inducing

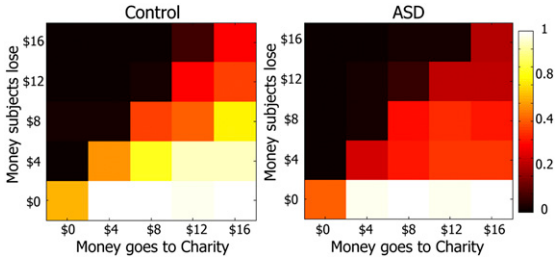

Fig. 2. Proportions of accepted donations in the Absence condition in each design cell depicted in Fig. 1*A*. White indicates a high proportion of accepted donations, whereas black indicates a low proportion.

more donations in the control group [$t(10) = 2.29$, $P = 0.02$], the ASD group showed a nonsignificant decrease in donation in the Presence condition (Fig. 3*A*).

In addition to effects of an observer on donations, as we found in the control participants, one would also expect an interaction between the strength of this observer effect and each individual participant's propensity to donate in the first place. Plausibly, those inclined already to donate a lot of money even in the absence of an observer would not donate much more when observed, whereas those donating little or nothing when alone might feel more motivated to donate and improve their reputation when observed. We found, in the control group, a highly significant negative correlation between the number of accepted donations in the Absence condition and the strength of the observer effect (difference in the number of accepted donations between the Presence and Absence conditions) ($r = -0.88$, $P < 0.001$), whereas there was no significant correlation in the ASD group ($r = -0.31$, $P = 0.38$, n.s.) (Fig. 3*B*); the difference between two correlation coefficients (controls vs. ASD) was also significant ($P < 0.05$, two-tailed). Although subjects could accept a theoretical maximum of 48 donation decisions per session, few accepted donations in which they had to pay more than the charity would gain (Fig. 1*A*, blue and cyan cells), a rational behavior because they could personally donate the same amount of money with less cost after the experiment. There is thus likely to be a psychological ceiling at 28 donation decisions, whereby personal loss and charity gain are equal (Fig. 1*A*, red plus yellow plus green cells). If we exclude those subjects who accepted 28 times or more in the Absence condition, as can be seen in Fig. 3*B*, six out of seven control subjects showed the observer effect (more donations in the Presence condition), whereas only one out of eight ASD subjects showed the effect; this difference in proportion was significant (Fisher exact test, $P = 0.009$). To confirm the robustness of the difference in observer effects between two groups, we ran a linear regression with the difference in accepted donations (Presence minus Absence conditions) as the dependent variable and included as regressors age, IQ, sex, and the number of accepted donations in the Absence condition. We replicated a significant effect of subject group ($\beta = 3.71$, $P = 0.001$) as well as the number of accepted donations in the Absence condition ($\beta = -0.20$, $P = 0.003$; all other $P > 0.24$), confirming that age, IQ, sex, and the intrinsic motivation to donate (i.e., the number of donations without observation) cannot account for our reported differences in observer effects between subject groups. Finally, to control for the possibility that ASD subjects might simply be perseverating in their decisions, we investigated possible group differences in the tendency to switch choices for a given design cell (Fig. 1*A*) across the two conditions, but we found no significant difference in the mean number of times they changed their choices (controls = 2.73 vs. ASD = 2.30; Mann-Whitney *U* test, $P = 0.25$, n.s.). Thus, the tendency to repeat the same choices across the two conditions did not differ between two groups.

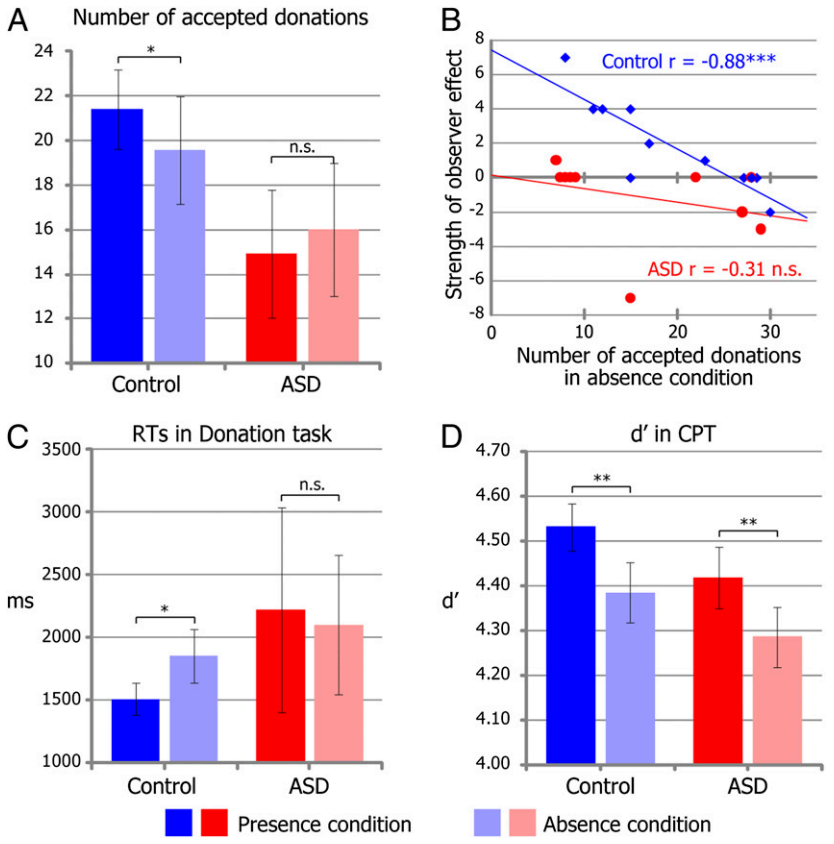

Fig. 3. Results for Donation and CPT tasks. Blue indicates control subjects, and red indicates ASD subjects. Dark blue/red indicates the Presence condition, and light blue/red indicates the Absence condition. (*A*) Mean number of accepted donations in each Presence and Absence condition for both groups. (*B*) Correlations between the number of accepted donations in the Absence condition and the susceptibility to the observer effect (difference in accepted donations between Presence vs. Absence condition). Higher value in the *y* axis indicates more donations in the Presence condition relative to the Absence condition. Values of the *x* axis are jittered to reduce the overlap of points. (*C*) Mean RTs in the Donation task. (*D*) Average d' in CPT. Higher d' means higher sensitivity to target stimuli. For *A*, *C*, and *D*, *P* values were based on one-tailed paired *t* tests. Error bars indicate SEM. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

Reaction Times. Reaction time (RT) data in the Donation task also showed an effect of the Observer condition in the control but not ASD group (Fig. 3*C*). To control for the effect of task familiarity on RTs, we included the order of the two sessions (Presence session first or Absence session first) as another between-subject factor. A 2 (group) × 2 (observer) × 2 (session order) mixed ANOVA showed a trend effect for a group × observer interaction [$F(1,17) = 3.75$, $P = 0.070$] as well as a significant observer × order interaction [$F(1,17) = 7.89$, $P = 0.012$]. No other effect was significant (all $P > 0.22$). As a follow-up, we ran within each subject group a 2 (observer) × 2 (order of session) mixed ANOVA, which revealed main effects of observer ($P = 0.006$) and session order ($P = 0.008$) as well as their interaction ($P = 0.036$) in the control group, but no significant effects in the ASD group (all $P > 0.12$). These findings suggest that the group differences in observer effects we reported earlier are, to some extent, also reflected in RT data.

Continuous Performance Task. We also had participants carry out a continuous performance task (CPT) in the presence or absence of an observer, to determine whether the observer effects we reported above for the donation task truly reflect differential effects of social reputation or a broader deficit in social cognition in the ASD group (such as an inability even to represent the presence of another person). For the CPT task, both ASD and control subjects were highly accurate in detecting target stimuli (99.4% and 99.6%, respectively), and there was no difference in overall accuracy. We calculated d' as the dependent variable for each subject and ran a 2 (group) × 2 (observer) × 2 (session order) mixed ANOVA. We found only a significant main effect of observer [$F(1,17) = 16.7$, $P = 0.001$], indicating that for both ASD and control groups their performances were better in the presence of an observer than when alone (Fig. 3*D*). The same mixed ANOVA on response bias revealed no significant effect (all $P > 0.28$). Furthermore, the mixed ANOVA on RTs during the CPT revealed only a significant main effect of session order [$F(1,17) = 7.10$, $P = 0.016$], indicating that RTs of those who did the Presence session first were faster than those who did the Absence session first, regardless of group.

Questionnaire-Based Measures. The effect of an observer's presence on mood was assessed with the Positive and Negative Affect Schedule (PANAS) (25), a standardized questionnaire assessing current positive and negative moods. A 2 (group) × 2 (observer) mixed ANOVA (separately for positive and negative affect) revealed no significant effects on either positive or negative affect (all $P > 0.28$). Furthermore, within each group, neither positive nor negative mood were correlated with the number of accepted donations in each condition (all $P > 0.26$).

We also administered a postexperiment questionnaire that provided further personality-related measures (*Materials and Methods*). Mean ratings on the Social Desirability scale (26), a measure of the need for social approval, were no different between two groups ($P = 0.53$, two-tailed). Although a prior study has suggested that individuals scoring higher in their need for social approval were also more susceptible to observer effects during prosocial decision making (5), we found no correlation with the strength of the observer effect on our Donation task in either subject group (control $r = 0.10$, n.s., and ASD $r = 0.18$, n.s.).

We also asked questions measuring attitude toward the charity we used [United Nations Children's Fund (UNICEF)] and their perception of the social desirability of donating to this charity. Subject groups did not differ in their attitude (control mean = 5.27 vs. ASD mean = 4.55; $P = 0.36$, two-tailed) or their perception of social desirability of donating (control mean = 4.55 vs. ASD mean = 4.90; $P = 0.62$, two-tailed).

Quantifying Observer Behavior. To verify that there was no difference between subject groups in the behavior of the experimenter who was acting as the observer in our study, independent raters analyzed video recordings that were made covertly during the Presence session. Coding of these tapes by two independent coders (who were blind to the group membership of the subject) confirmed that there was no occasion on which the observer engaged differentially in any apparent activities (e.g., talking, coughing, etc). Additionally, after checking each videotape, two coders were encouraged to guess whether the observer was watching ASD or control participants; their best guesses were at chance (Fisher exact test, all $P > 0.67$), indicating that there was no detectable difference in the observer's behavior between the two groups.

Discussion

The present study showed that whereas control subjects donated more often in the presence of an observer than when they made donation decisions alone, ASD subjects showed no such effect (if anything, a slight trend in the opposite direction). Moreover, there was a correlation in the controls between how much they were inclined to donate without observation and the strength of the observer effect; and there was an effect on RT due to the presence of the observer. None of these effects were present in people with ASD. The equivalent social facilitation effects seen in both groups on a CPT task argue that people with ASD have intact nonspecific effects of the presence of another person and can perceive other people. Taken together, the findings indicate that individuals with ASD have a specific deficit in taking into account their reputation in the eyes of others.

Might people with ASD be immune to observer effects simply because they have less empathy for others (less intrinsic motivation to help others)? We did find that ASD features somewhat fewer donations in general, in addition to the more specific lack of an observer effect on these donations (Fig. 3*A*), even though this was not a statistically significant difference between groups. It has been well established that ASD features reduced empathy (27, 28), and our ASD sample actually scored significantly lower on Simon Baron-Cohen's Empathy Quotient than the control group, whereas there was no difference in Systematizing Quotient (29) between the two groups (Table S1). However, in the present study, by keeping the partner of the game identical (UNICEF) and only manipulating the presence of an observer (thus manipulating only the extrinsic motivation to donate), any overall differences in empathy (intrinsic motivation to donate) cannot explain the group differences on the observer effect that we report.

Further evidence for the specificity of the impairment in social reputation processing that we report in ASD comes from the CPT task. ASD subjects showed a normal effect of the presence of another person on this task (Fig. 3*D*), likely indicating a normal ability of the presence of others to induce increased arousal or social facilitation (24). This finding also argues, at least in our high-functioning group of people with ASD, that basic perception of the other person was normal, as further confirmed by normal performances on the Benton Facial Recognition Task (30) (Table S1). Finally, there were no group differences in mood, need for approval, attitude toward the charity, or social desirability from postexperiment questionnaires. We suggest that ASD features a specific impairment in representing social reputation—the esteem in which others hold us—and in linking social reputation to the motivation for public prosocial behaviors.

It is intriguing that we found a strong negative correlation in the control group between the strength of the observer effect and the number of accepted donations in the Absence condition (Fig. 3*B*): the less likely healthy people are intrinsically to donate (in the absence of an observer), the greater their increase in donations upon being observed. A natural explanation of this pattern is that it is those participants who donate least who stand to improve their reputation the most by changing their behavior in the eyes of others. The lack of such a correlation in the ASD group provides further evidence for a lack of reputation-based processing. It is worth noting also that the negative correlation we found in the controls together with the somewhat lower overall donations we found in the ASD group also suggest that, if anything, the group differences in observer effects we found may have been underestimated.

The RTs of controls, but not ASD participants, were on average approximately 400 ms faster in the presence of an observer (Fig. 3*C*). It is unlikely that this effect was due merely to an increase in attention or arousal in the observer's presence because we should have found that effect also in the ASD group, given their normal social facilitation effects on the CPT task. Instead, we find it plausible that the observer effects on donation RTs seen in the controls were also driven by social reputation: the faster the decision to donate something, the better the reputation.

What mechanisms might account for ASD individuals' insensitivity to social reputation? Prior studies have demonstrated that social reputation can provide a powerful incentive for prosocial behavior (31) and that the underlying mechanism may recruit general reward-processing regions of the brain (32, 33). That is, in healthy individuals, improving one's social reputation acts as an instrumental reinforcer because better social reputation is rewarding. We think that there are at least two possible explanations for this deficit in ASD individuals. The first possibility is that they can represent the presence of an observer but may be unable to take the extra metacognitive step of representing what the observer thinks of them (reputation). The second possibility is that they can represent the observer as well as their reputation but lack normal social reward processing. That is, social reputation may not be rewarding and would thus fail to influence their behavior in our task. Past reports on ASD individuals' difficulty in representing the mental states of others (17, 18) suggest that they may lack the metacognitive ability to understand the reputation they have with others (10, 23, 34) and thus favor the first explanation. However, there are also findings that although people with high-functioning ASD can attribute mental states to others if explicitly asked to, they fail to do so spontaneously (35), suggesting that there may be a primary motivational deficit. Consistent with this idea is a recent finding that stimuli that are normally social rewarding (smiling faces) fail to activate reward circuitry in children with autism (36). Future studies will be required to disentangle precisely at which stage of processing the deficit occurs that we report here (see below for a possible idea).

The present results demonstrate that prosocial behavior in ASD is insensitive to the effects of an observer, supporting the hypothesis that ASD features impaired processing of social reputation. This may well account for some of the real-world social deficits of ASD, but there remain several important topics for future investigation. First, it will be important to extend the present findings to other circumstances encountered in everyday life. Although our study focused on the good side of the observer effect (increased prosocial behavior), there is also its dark side: one sometimes feels more anonymous in a large crowd (exhibiting less concern for reputation). The presence of many other people could therefore lead to less prosocial performance (e.g., social loafing; ref. 37) or to increased antisocial behavior (e.g., deindividuation; ref. 38). Testing these phenomena in people with ASD could provide additional evidence for their insensitivity to the presence of other people.

Relatedly, it will be important to link the present findings from a somewhat contrived situation in the laboratory to real-world clinical relevance. Do people with ASD evidence insensitivity to the presence of other people in real-world contexts? Additionally, are such deficits mediated by impaired social reputation processing? The present results support such a hypothesis, but additional studies that carefully characterize actual real-world

behavior will be required to definitively establish this link. Plausibly, high-functioning people with ASD will show impaired social reputation effects under some circumstances (such as those in our experiment) but not others (such as those providing additional explicit and contextual cues on the basis of which compensatory processing might take place). We thus take the present findings as only the first step and acknowledge that impairments in social reputation processing may well be heterogeneous in ASD and may be partly compensated in high-functioning individuals if more explicit cues are present.

A second important topic for future investigation is to characterize the behavior of people with ASD in other economic games that provide the opportunity for reputation. As we noted earlier, prior studies using trust games (19) and prisoner's dilemma games (20, 21) have failed to show a clear difference between ASD and control groups, but subjects in those studies played multiple rounds of a game with the same partner. Repeated matching with a partner involves "direct" reciprocity, fulfilling a personal obligation to a specific individual. This kind of interaction does not require salient judgments of how one's behavior influences one's social reputation in the eyes of a more distant observer or in the eyes of people generally. Because our present results highlight an impairment in sensitivity to social reputations that noninteractive observers hold, people with ASD might well perform abnormally in repeated cooperation games in large groups in which players are matched randomly with one another. Studies in healthy populations indicate that in games relying on single interactions with a partner, people do behave more prosocially, provided that a reputation system guarantees that their behavioral history becomes known to future partners (an instance of "indirect reciprocity") (39). In a game like this, the primary motivation for players to act prosocially may not be good reputation per se but money into which good reputation is converted (i.e., players with good reputations usually get more cooperation from others and end up making more money at the end of the experiment). Therefore, as long as subjects possess the ability to form a metarepresentation of how one is generally viewed by possible future players (and money is rewarding to them), individuals cooperate even if a good reputation is per se not rewarding. Thus, by comparing how participants with ASD might behave in such a game with vs. without a reputation system, it might be possible to tell whether their deficit is primarily cognitive or motivational.

Third, it will be important in future studies to further delineate the neural underpinnings of the behavioral effect we report here, which might provide considerable insight into its underlying mechanism. Prior studies have pointed to structures such as the cingulate cortex in representing one's own intentions (40) and have documented specifically abnormal cingulate activation corresponding to such self-relevant processing in autism (19). Other candidate structures that may contribute to the impaired social reputation processing of ASD include parts of medial and polar prefrontal cortex involved in metacognitive representation (10), which have been implicated in reputation processing in healthy individuals (33, 41), as well as white matter connectivity that serves to link the various representations required and that is likely to be impaired in autism (42).

Materials and Methods

Subjects. Twenty-four subjects participated in the present study. The reported analyses were based on 21 subjects: 10 subjects with ASD (2 female) and 11 age- and IQ-matched controls (1 female; Table S1). One ASD and 2 control subjects were excluded from the analysis because they repeatedly rejected choices on the trials that involved no cost to themselves (there are 16 such trials in total across two sessions; Fig. 1*A*, red cells), raising the concern that they did not understand the task correctly. All ASD participants met the *Diagnostic and Statistical Manual of Mental Disorders, Revised 4th Edition* diagnostic criteria for autism or Asperger syndrome and met the cutoff scores for autism or Asperger syndrome on the Autism Diagnostic Observation Schedule, Module 4 (43). We could also collect the Autism Diagnostic Interview–Revised (44) from seven ASD subjects, and all seven of them met the cutoff scores on this scale (Table S1). Participants gave informed consent to participate in the studies under a protocol approved by the Institutional Review Board of the California Institute of Technology.

Experimental Tasks. The experiment consisted of two sessions. In each session, subjects performed a Donation task and a CPT in counterbalanced order, and they also answered a mood questionnaire (PANAS) at the end; additional questionnaires were administered at the very end of the experiment (see below).

Our primary dependent measure of interest came from performance on a Donation task (Fig. 1 *A* and *B*), in which subjects performed a dictator game in which the amount of money subjects themselves lost and the amount of money gained by the charity were independently manipulated (range \$0 to \$16 by an increment of \$4; Fig. 1*A*). In each trial, participants had to accept or reject a monetary transfer presented on the screen. To minimize the effect of memory (simply repeating the same choices between Presence and Absence sessions), the actual amount of money presented in each trial during the experiment was jittered (Fig. 1*B*) by adding a random number drawn from a normal distribution with a mean and SD of 0 and 0.3, respectively. Jittering was not added if the original amount was \$0. Only for diagonal cells (green cells in Fig. 1*A*), exactly the same amount of jittering was added to both the money subjects lose and the money the charity gains, so that in these trials the amount of money subjects lose was always equal to the amount of money the charity gains. Each design cell was implemented twice in each of the Presence and Absence sessions (i.e., 50 trials each) in randomized order. After each decision, chosen options ("Accept" or "Reject") were highlighted by a red circle, to ensure that choices were clearly observable by the observer in the Presence condition. All subjects were given \$45 as an endowment at the beginning of the experiment and told that one trial would be randomly chosen at the end of the experiment, and their decision on that trial was implemented (including real donations to UNICEF through the Internet).

As a comparison task, we used a CPT that engaged multiple cognitive processes and required sustained attention and vigilance but did not have any social aspect (45). During the task, subjects were presented with letters one at a time on the screen and were asked to press a key every time they saw the letter "X." Each letter was presented for 200 ms (1,000-ms interstimulus interval). The total duration of the test was ≈13 min (646 trials in total), and the target letter ("X") was presented in 26.5% of all trials. For both Donation and CPT tasks, stimuli were presented using MATLAB with Psychtoolbox.

Experimental Procedure and Observer Manipulation. In one session, subjects' performances were constantly observed by an observer, whereas in the other session they completed the tasks alone. At the beginning of the experiment a male experimenter gave instructions for both the Donation and CPT tasks and provided information about the mission of UNICEF and how the money donated will be used. Next, subjects performed a brief practice session for both tasks before the actual experiment. During the instruction, the experimenter was blind to the upcoming order of sessions (Presence or Absence session first), of which he was informed during the practice session. If the first session was the Absence session, the experimenter left the room after he set up the presentation program, and subjects performed the Donation and CPT tasks and completed a PANAS questionnaire alone in the room.

If the first session was the Presence session, the experimenter set up a task presentation program and left the room. However, the program in this condition was programmed to simulate a crash after approximately 1 to 2 min: after 8 donation trials (if they performed Donation task first) or 28 CPT trials (if they performed CPT first), the task screen suddenly disappeared, and subjects saw error messages written in red font on the MATLAB command window and heard a beep sound. When this occurred, all subjects except one ASD subject spontaneously came out of the room and reported to the experimenter that the task had crashed. For the single ASD subject who did not come out, the experimenter entered the room ≈5 min after he had left and asked the subject if everything was fine; the ASD participant reported that he was about to go out. In each case, the experimenter apologized for the malfunction and asked subjects to wait in a different room while, ostensibly, he was fixing the task program.

After ≈5 min of waiting, subjects were asked to come back to the experimental room, and the experimenter briefly introduced an unfamiliar male research assistant they had never met before (a confederate who played the role of observer). Subjects were told that because it was not certain that the program was completely fixed and that all data would be correctly saved, this technician would stay in the room with them and watch and write down their choices during the Donation task just to be sure that the data were recorded (subjects were also told that the observer would not record their performance during the CPT, but stay there in case the program crashed again). The observer quietly sat ≈3 feet diagonally behind the subject throughout the session. Although subjects had a vague sense of the observer behind them, the computer monitor they were facing was not

glossy, and they could not see the reflected observer's face or small body motions during the experiment. To verify that there was no difference in the observer's behavior between the two subject groups, the observer was videotaped through a one-way mirror by a video camera placed in the next room (the one-way mirror was mostly covered by a blind and a white board, rendering it completely inconspicuous). While being observed by the observer, subjects completed the two tasks and PANAS. When they finished all tasks, the observer thanked subjects and left the room to inform the experimenter.

After completing the first session, each subject participated in a variety of other experiments in our laboratory that were part of different ongoing studies in autism research (e.g., answering personality questionnaires, preference judgments of various stimuli, etc). There was no systematic difference in what intervening tasks were performed between the two subject groups. After completing these other experiments, subjects took a brief break and then started the second session. Therefore, the two sessions of the present experiment were separated by 50–70 min.

At the beginning of the second session, the experimenter briefly instructed subjects to complete the same two tasks again. If the second session was the Absence session, subjects were told that the program had been thoroughly fixed, and they completed the tasks and PANAS alone in the room. On the other hand, if the second session was the Presence session, the above-mentioned procedure was followed (i.e., the presentation program crashed, and an observer was introduced), and subjects competed the tasks in the presence of an observer.

The order of the two sessions (Presence or Absence sessions) and the order of the two tasks (Donation task or CPT) was counterbalanced across subjects, and within each subject the order of the two tasks was fixed across the two sessions.

After completing the second session, subjects moved to a different room, and they answered the Social Desirability Scale (26), which measures people's tendency to respond in a socially desirable manner, and two follow-up questions: "To what extent do you think the mission of UNICEF is important?" and "To what extent do you think that making a donation to UNICEF is socially desirable?" The first question measured subjects' personal attitude toward the charity, and the second one was intended to measure their perception of how the charity is valued by a society or other people. All of these questions were answered using a 7-point scale. Subjects were also asked what they thought was the purpose of the experiment they had just completed. No subject mentioned any purpose involving effects of an observer, and no subject thought that the crash of the program was intentional. Finally, one trial was selected and their choice on that trial was implemented.

Analysis. For Donation task data, we excluded two 0–0 trials (Fig. 1*A*, gray cell) in each session from the analysis and analyzed donation decisions and RTs for the remaining 48 trials. For CPT data, we computed d' and response bias separately for each session for each subject. RTs in correct trials were also analyzed. For RT analysis for both tasks, within each individual subject, RTs deviating from the subject's mean by more than 3SDs were excluded from further analysis.

Because of the a priori expected direction of the observer effects (i.e., better performance in CPT, more "Accept" in the Donation task in the Presence session), one-tailed *P* values are reported (unless otherwise specified) when comparing the Presence vs. Absence sessions.

ACKNOWLEDGMENTS. We thank Brian Cheng and Catherine Holcomb for assistance in recruiting subjects and conducting the experiments, Tim Armstrong and Margaret Lee for coding videotapes, and Drs. Lynn Paul and Dan Kennedy for help with diagnoses and assessments of the participants. This study was supported by a Grant-in-Aid for Japan Society for the Promotion of Science Fellows (to K.I.), grants from the Simons Foundation and the National Institute of Mental Health (to R.A.), and the Tamagawa University Global Centers of Excellence grant from Ministry of Education, Culture, Sports, Science and Technology, Japan.

1. Leary MR, Kowalski RM (1990) Impression management: A literature review and two-component model. *Psychol Bull* 107:34–47.
2. Camerer CF (2003) *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton Univ Press, Princeton).
3. Bereczkei T, Birkas B, Kerekes Z (2007) Public charity offer as a proximate factor of evolved reputation-building strategy: An experimental analysis of a real-life situation. *Evol Hum Behav* 28:277–284.
4. Kurzban R, DeScioli P, O'Brien E (2007) Audience effects on moralistic punishment. *Evol Hum Behav* 28:75–84.
5. Satow KL (1975) Social approval and helping. *J Exp Soc Psychol* 11:501–509.
6. Bateson M, Nettle D, Roberts G (2006) Cues of being watched enhance cooperation in a real-world setting. *Biol Lett* 2:412–414.
7. Haley KJ, Fessler DMT (2005) Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol Hum Behav* 26:245–256.
8. Warneken F, Hare B, Melis AP, Hanus D, Tomasello M (2007) Spontaneous altruism by chimpanzees and young children. *PLoS Biol* 5:e184.
9. Call J, Tomasello M (2008) Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn Sci* 12:187–192.
10. Amodio DM, Frith CD (2006) Meeting of minds: The medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277.
11. Alexander R (1987) *The Biology of Moral Systems* (Aidine de Gruyter, New York).
12. Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393:573–577.
13. Association AP (APA) (2000) *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, Washington, DC), revised 4th Ed.
14. Frith U (2003) *Autism: Explaining the Enigma* (Blackwell, Oxford), 2nd Ed.
15. Losh M, et al. (2009) Neuropsychological profile of autism and the broad autism phenotype. *Arch Gen Psychiatry* 66:518–526.
16. Pelphrey K, Adolphs R, Morris JP (2004) Neuroanatomical substrates of social cognition dysfunction in autism. *Ment Retard Dev Disabil Res Rev* 10:259–271.
17. Frith U (2001) Mind blindness and the brain in autism. *Neuron* 32:969–979.
18. Baron-Cohen S (1995) *Mindblindness: An Essay on Autism and Theory of Mind* (MIT Press, Cambridge, MA).
19. Chiu PH, et al. (2008) Self responses along cingulate cortex reveal quantitative neural phenotype for high-functioning autism. *Neuron* 57:463–473.
20. Downs A, Smith T (2004) Emotional understanding, cooperation, and social behavior in high-functioning children with autism. *J Autism Dev Disord* 34:625–635.
21. Hill EL, Sally D, Frith U (2004) Does mentalising ability influence cooperative decision-making in a social dilemma? Introspective evidence from a study of adults with autism spectrum disorder. *J Conscious Stud* 11:144–161.
22. Yoshida W, et al. (2010) Cooperation and heterogeneity of the autistic mind. *J Neurosci* 30:8815–8818.
23. Frith CD, Frith U (2008) The self and its reputation in autism. *Neuron* 57:331–332.
24. Zajonc RB (1965) Social facilitation. *Science* 149:269–274.
25. Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. *J Pers Soc Psychol* 54:1063–1070.
26. Crowne DP, Marlowe D (1960) A new scale of social desirability independent of psychopathology. *J Consult Psychol* 24:349–354.
27. Baron-Cohen S, Belmonte MK (2005) Autism: A window onto the development of the social and the analytic brain. *Annu Rev Neurosci* 28:109–126.
28. Baron-Cohen S, Wheelwright S (2004) The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J Autism Dev Disord* 34:163–175.
29. Baron-Cohen S, Richler J, Bisarya D, Gurunathan N, Wheelwright S (2003) The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philos Trans R Soc Lond B Biol Sci* 358:361–374.
30. Benton AL, Sivan AB, Hamaher K, Varney NR, Spreen O (1994) *Contributions to Neuropsychological Assessment* (Oxford Univ Press, New York).
31. Benabou R, Tirole J (2006) Incentives and prosocial behavior. *Am Econ Rev* 96: 1652–1678.
32. Izuma K, Saito DN, Sadato N (2010) Processing of the incentive for social approval in the ventral striatum during charitable donation. *J Cogn Neurosci* 22:621–631.
33. Izuma K, Saito DN, Sadato N (2010) The roles of the medial prefrontal cortex and striatum in reputation processing. *Soc Neurosci* 5:133–147.
34. Frith U, Frith C (2010) The social brain: Allowing humans to boldly go where no other species has been. *Philos Trans R Soc Lond B Biol Sci* 365:165–176.
35. Senju A, Southgate V, White S, Frith U (2009) Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science* 325:883–885.
36. Scott-Van Zeeland AA, Dapretto M, Ghahremani DG, Poldrack RA, Bookheimer SY (2010) Reward processing in autism. *Autism Res* 3:53–67.
37. Latane B, Williams K, Harkins S (1979) Many hands make light the work—causes and consequences of social loafing. *J Pers Soc Psychol* 37:822–832.
38. Festinger L (1952) Some consequences of de-individuation in a group. *J Abnorm Soc Psychol* 47:382–389.
39. Engelmann D, Fischbacher U (2009) Indirect reciprocity and strategic reputation building in an experimental helping game. *Games Econ Behav* 67:399–407.
40. Tomlin D, et al. (2006) Agent-specific responses in the cingulate cortex during economic exchanges. *Science* 312:1047–1050.
41. Izuma K, Saito DN, Sadato N (2008) Processing of social and monetary rewards in the human striatum. *Neuron* 58:284–294.
42. Geschwind DH, Levitt P (2007) Autism spectrum disorders: Developmental disconnection syndromes. *Curr Opin Neurobiol* 17:103–111.
43. Lord C, et al. (2000) The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord* 30:205–223.
44. Lord C, Rutter M, Le Couteur A (1994) Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 24:659–685.
45. Beck LH, Bransome ED, Jr., Mirsky AF, Rosvold HE, Sarason I (1956) A continuous performance test of brain damage. *J Consult Psychol* 20:343–350.