

Get them before they get you: Trust, trustworthiness, and social cognition in boys with and without externalizing behavior problems

CARLA SHARP,^a CAROLYN HA,^a AND PETER FONAGY^b

^aUniversity of Houston; and ^bUniversity College London

Abstract

Economic exchange games have rarely been applied to examine psychopathology in youth. In the current study we adapted a trust game to investigate the relations between externalizing behavior problems, trust, and trustworthiness. We were particularly interested in the differential modulating impact of “known identity” (vs. anonymous) condition of the task. Second, we examined whether anomalies in trust behavior would correspond to social cognition manifested in children with externalizing problems. A total of 171 (79 age-matched pairs) boys (mean age = 12.84; *SD* = 1.80) were recruited from community groups where social networks and relationships amongst peers have been established. A trust game was played under two conditions: an anonymous version where the identity of the trust game partner was not known and a “known identity” version where identities were revealed. Results supported the conclusion that boys with externalizing behavior problems are generally less trustworthy, but not less trusting and that this was true especially for the known identity version of the game. Moreover, anomalies in trust behavior were associated with hostile intentions, but not reflective of a general theory of mind deficit. This study contributes to an emerging literature using economic exchange games to investigate real-time, real-life exchanges in relation to psychopathology.

Externalizing behavior problems refer to a broad range of disruptive antisocial behaviors as captured by the diagnoses of conduct disorder and oppositional defiant disorder (American Psychiatric Association, 2000). One of the hallmark features of externalizing problems is interpersonal difficulties. Children with externalizing problems tend to have poor relationships with peers (Vitaro, Tremblay, & Bukowski, 2001) and parents (Greenberg, Speltz, DeKlyen, & Endriga, 1991). Social-cognitive theory provides one framework for understanding the relation between externalizing problems and interpersonal difficulties. Research examining social cognition in children with externalizing behavior problems has been dominated by what is now known as the social information processing (SIP) model of aggression (Dodge, 1993; Dodge et al., 2003). In this model, biases and deficits in several sequential cognitive processing steps are thought to be the proximal mechanisms for aggressive interpersonal behavior in specific social situations (Mize & Pettit, 2008). Other social-cognitive constructs that have been shown to associate with externalizing behavior problems include theory of mind (e.g., Hughes & Ensor, 2008; Sharp, 2008), distorted mentalizing (e.g., Fo-

nagy et al., 2009; Sharp, Croudace, & Goodyer, 2007), emotional empathy (e.g., Lovett & Sheffield, 2007) and perspective taking (e.g., Dodge, 1980).

Although the impact of the SIP model and other approaches has been substantial, there are some limitations to previous approaches. For instance, although associations between hostile interpretations of hypothetical vignettes and a history of aggressive behavior are replicable, the magnitude of observed associations is small (Dodge & Somberg, 1987; Orobio de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002). Small effect sizes may be due to the “off-line” nature of the tasks by virtue of reliance on hypothetical scenarios. Tasks are therefore not administered in real time, they do not sample actual social interactions, and they are unlikely to elicit full emotional and behavioral engagement. Most social-cognitive tasks are also characterized by an overreliance on self-report, thereby “pulling for thinking” (Mize & Ladd, 1988; Mize & Pettit, 2008) and eliciting socially desirable responses. More ecologically valid approaches are needed to build on existing social-cognitive research of psychological disorders in general, and externalizing problems in particular.

Experimental tasks from the economic exchange literature (for a review, see Camerer, 2003) offer a new approach to operationalizing social cognition in real-life, emotionally charged interaction in controlled settings (McClure et al., 2007). One such experimental task, the trust task, is based on an evolutionary model that draws on game theory principles in which trust is defined as a tit-for-tat social exchange (reciprocal altruism; Axelrod, 1986). The trust task was initially proposed by Camerer

This research was funded by a Competitive Seed Funding Program at Baylor College of Medicine. We thank the adolescents and their families who participated in the research, and we are grateful to Amanda Venta and Tyson Reuter for reviewing the manuscript.

Address correspondence and reprint requests to: Carla Sharp, Department of Psychology, 126 Heyne Building, University of Houston, Houston, TX 77024; E-mail: csharp2@uh.edu.

and Weigelt (1988) and further developed by Berg, Dickhaut, and McCabe (1995). One player (the investor) is endowed with a certain amount of money (or points as proxies for money). The investor can keep all the money or decide to “invest” some amount with the partner (the trustee). The amount invested is tripled in value as it is sent to the trustee, who then decides what portion to return to the investor. Research has shown that the investor typically makes a substantial offer to the trustee so that the offer is perceived as fair (Camerer, 2003). In turn, the trustee responds to the initial offer from the investor with a reasonable relative return offer. In so doing, the investor displays “trust” (making himself vulnerable by taking a course of action that creates incentives for the other party to exploit him) while the trustee displays “trustworthiness” (reciprocity; Ashraf, Bohnet, & Piankov, 2006).

Although social-cognitive processes are not directly measured in these games, it is assumed that trust behavior relies on social-cognitive capacity. In particular, it has been suggested that trust behavior requires the capacity to detect or predict the intentions of the other player (Dufwenberg & Gneezy, 2000; Falk & Fischbacher, 2006; McCabe, Smith, & LePore, 2000). Trust behavior also requires the capacity to view the game from the other player’s perspective (Singer & Fehr, 2005). Recently, Van den Bos, Westenberg, Van Dijk, and Crone, *in press*) examined the role of perspective taking in trust game behavior by defining perspective taking as the ability to consider the intentions of others as well as the consequences of subjects’ own trust game behavior for others. By varying these outcomes (risk vs. benefit) across different age groups (9–25 years), the authors demonstrated age-related changes in sensitivity to outcome for partners as indexed by subjects’ behavioral choices. Specifically, they found age-related changes in sensitivity to the benefit of the other player in trust decisions. Although perspective taking was not directly measured, the behavioral results suggested that as social perspective taking capacity increased with age, so did trust behavior.

There are several reasons for believing that trust behavior, and the social-cognitive reasoning accompanying trust behavior, may be anomalous for children with externalizing behavior problems. First, as summarized above, links have already been established between externalizing behavior problems and perspective taking (Dodge, 1980), theory of mind (e.g., Hughes & Ensor, 2008; Sharp, 2008), and mentalizing (e.g., Sharp et al., 2007). Moreover, children with externalizing problems show hostile attribution biases in interpreting the intent of others in ambiguous social situations (Mize & Pettit, 2008; Orobio de Castro et al., 2002). If intentionality detection and perspective-taking abilities underlie trust behavior as suggested by Van den Bos et al. (*in press*), Singer and Fehr (2005), McCabe et al. (2000), and others (Pillutla, Malhotra, & Murnighan, 2003; Malhotra, 2004), it follows that children with externalizing behavior problems may be expected to show anomalous trust behavior. An important additional factor that may influence trust behavior is the identity of trust game partners. The identity of target characters in hypothetical

scenarios has been shown to have powerful effects on children’s assumptions about target characters’ motives and behaviors (Mize & Pettit, 2008). Dodge (1980) showed that children were five times more likely to attribute hostile intentions to another child if that child was known to be aggressive.

Against this background, the current study had two aims. First, we aimed to test whether there were differences in trust and trustworthiness between boys with externalizing behavior problems and those without, and whether knowing the identity of game partners had differential effects on boys with and without externalizing problems. Based on established links between perspective-taking difficulties (e.g., Dodge, 1980), empathizing difficulties (Lovett & Sheffield, 2007), theory of mind (Sharp, 2008), and externalizing behavior disorders, we expected lower trust and trustworthiness in boys with externalizing problems. Based on research showing that empathy is increased during the trust game when the identity of partners are known (Camerer, 2003; Jenni & Loewenstein, 1997), we also predicted that anonymity would attenuate trust and trustworthiness such that higher trust and trustworthiness will be displayed when the identity of trust game partners were known to subjects. We expected that this effect would be smaller for boys with externalizing behavior problems, because of reduced empathy associated with externalizing problems (Lovett & Sheffield, 2007).

Our second aim was to test whether social-cognitive capacity served as a proximal mechanism of trust behavior. This assumption has not yet been directly tested. It has become increasingly clear that social cognition is a multicomponent construct (Sharp, Fonagy, & Goodyer, 2008). Therefore, we used two measures to assess social cognition. First, off-line theory of mind capacity was assessed using a standard and advanced test of theory of mind (the Child’s Eye Task; CET; Baron Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), which has been shown to associate with conduct problems in preadolescent children (Sharp, 2008). The CET measures individuals’ capacity to read the mental states of others in a standard set of stimuli of the eye region of the face only. Given the “off-line” nature of this task, we did not expect performance on this task to correlate with trust behavior. Second, the method for measuring social cognition involved assessment of subjects’ own intentions and their reading of the intentions of their partners during the trust task. Subjects were asked to judge the intentions behind the behavior of their partners (nonverbal cue detection) and to provide rationalizations (intentions) for their own decision making. Given the on-line nature of task-related social-cognitive reasoning, we expected correlations between the latter and trust behavior. More specifically, we expected that boys with externalizing behavior problems would report hostile (as opposed to benevolent) intentions underlying their own behavior and the behavior of others.

To reiterate, the current study had two aims. First, we examined the relationship between externalizing problems (self-reported, parent reported, and peer ascertained), trust and trustworthiness, in particular, the differential modulating impact of knowing the identity of trust game partners. Second,

we examined whether anomalies in trust behavior would correspond to hostile intentions reported by children with externalizing problems. Thus, the current study is the first to examine how trust and trustworthiness relate to externalizing behavior problems and social-cognitive reasoning in youth during real-life, real-time social interactions.

Methods

Participants

The study was approved by the relevant Institutional Review Board. A total sample of $n = 171$ of male youth between Grades 2 and 12 were recruited from community youth groups (Boy Scouts, 2010). Groups met weekly for activities and boys had known each other for an average of 2.3 years ($SD = 1.4$ years). Boy Scouts were purposefully recruited to exploit that these children have already established reputations (models) of each other. Recruitment procedures included first sending information about the study to parents. The research team's contact information was provided, and parents and children were encouraged to ask any questions they had about the study and/or consent procedures. Positive consent and assent were obtained from parents and children, respectively. Measures of externalizing behavior problems and demographic information were completed individually by parents and children. The collection of experimental data occurred during a single examination in groups of sizes between 6 and 15.

Thirteen boys did not have complete data on all measures and were excluded from analyses. They did not differ significantly from other subjects on any of the key study variables. The final data set consisted of 79 age-matched pairs of boys ($n = 158$). The mean age of the sample was 12.84 years ($SD = 1.80$) (seventh grade), with the youngest child aged 8 ($n = 1$) and the oldest aged 18 ($n = 1$). A total of 49.4% ($n = 78$) of children were between ages 8 and 12 and 50.6% ($n = 80$) of children were between ages 13 and 18.

Youths were matched only on age and were not screened for behavior problems before matching. Boys were randomly

assigned to the role of investor or trustee, after which trust tasks and self-report measures were concurrently administered. It was therefore entirely possible that boys above or below cutoff on measures of externalizing behavior problems could be over-represented in either the Investor or Trustee role. To ensure that this was not the case, we cross-tabulated the externalizing behavior variable with the trust game role variable. Although there were a higher number of boys in the above cutoff category for trustees ($n = 28$) compared to investors ($n = 19$), this was not significant ($\chi^2 = 0.03$, $df = 1$, $p = .86$). We also investigated age differences for boys above and below cutoff on externalizing behavior problems and found no difference ($t = 0.52$, $df = 156$, $p = .60$).

The sample was ethnically representative of the community from which they were drawn, approximately 62% Caucasians, 14% Hispanic, 11% African American, 10% Asian American, 2% Middle Eastern, and 1% represented other groups. Approximately 10% of sample's parents had obtained a high school diploma or equivalent certification, 18% received some college, 48% obtained a bachelors or associates degree, 17% received a masters degree, and 7% received doctorate level training. The socioeconomic status of the sample, based on parental educational attainment, was estimated to be primarily middle class.

Boy Scout samples, although technically "community" samples, may differ from normal samples because prosocial values are purposefully instilled in scout troops. Scout groups recruit youth equally that are at risk for behavior problems, that is, youth who are economically disadvantaged, live in single- or no-parent homes, and/or are growing up in hostile environments such as crime- or drug-infested neighborhoods ("Scoutreach," 2010). Although our sample was primarily middle class, our sample was comparable to community samples in terms of externalizing behavior problems. The ranges of scores on measures of externalizing behavior problems are presented in Table 1 along with comparison scores from community samples.

It is clear from Table 1 that the current sample was almost a perfect match to the Achenbach and Rescorla (2001) sample in terms of parent-reported externalizing problems. In terms

Table 1. Comparison with nonreferred sample on CBCL and YSR scores

	CBCL				YSR			
	Current Sample		Achenbach & Rescorla		Current Sample		Achenbach & Rescorla	
	Mean (<i>SD</i>)	Min–Max	Mean (<i>SD</i>)	Min–Max	Mean (<i>SD</i>)	Min–Max	Mean (<i>SD</i>)	Min–Max
Age	12.84 (1.8)	8–18	—	12–18	12.84 (1.8)	8–18	—	11–18
Total problems	49.73 (9.68)	24–77	49.80 (10.0)	—	55.55 (10.67)	32–81	50.1 (9.9)	—
Externalizing	48.71 (9.50)	33–80	50.1 (9.6)	—	53.83 (10.49)	29–79	50.2 (10.0)	—
Aggressive	53.86 (6.78)	50–94	54.3 (6.1)	—	57.08 (8.42)	50–93	54.4 (6.1)	—
Rule breaking	52.98 (4.69)	50–77	54.4 (5.8)	—	55.51 (7.15)	50–77	54.2 (5.7)	—
ODD	54.49 (6.09)	50–80	54.7 (5.8)	—	56.17 (6.78)	50–80	54.2 (5.5)	—
CD	53.49 (5.86)	50–84	54.2 (5.8)	—	57.17 (8.10)	50–84	54.3 (5.7)	—

Note: CBCL, Child Behavior Checklist (Achenbach & Rescorla, 2001); YSR, Youth Self-Report (Achenbach & Rescorla, 2001); ODD, oppositional-defiant disorder; CD, conduct disorder.

of self-reported externalizing problems, our sample showed slightly elevated levels of externalizing problems.

Experiments

Condition 1: Anonymous trust game procedure. To derive a measure of trust and trustworthiness, we adapted the basic, single-exchange trust game developed by Berg et al. (1995) in order to make it developmentally sensitive and appropriate for group administration (sizes 6–15) according to the procedure recommended by Sutter and Kocher (2007). Thus, the investor received 20 points of money ($X = 20$). The amount x sent by the investor was tripled as is typical for all trust games (Camerer, 2003). The trustee could then send back y , yielding final payoffs of $20 - x + y$ for the investor and $3x - y$ for the trustee.

The task was administered as a paper-and-pencil experiment where participants had to indicate their decisions on a decision form. The experiment began with the group of boys in the investor role. The game instructions were read out loud to investors, alongside written instructions. Investors were told that they were going to play a game with another boy (the trustee) from the same scout group whose identity would not be revealed. Investors were given the opportunity to ask questions about the instructions. None of the Investors asked for clarification indicating a good grasp of the instructions. Next, investors were asked to indicate on the decision form how many of 20 points they would like to send to the trustee. They were reminded that the number of points they sent would be tripled on the way to the trustee. Investors wrote down their decisions on the decision forms.

Verbal and written instructions were given to the trustees in a different room. Experimenters entered the investors' offers on trustees' decision forms. Trustees were reminded that the amount investors sent to them was tripled. Trustees then decided how much of the tripled amount they would like to send back to the investors. Decision forms were coded such that investor–trustee pairings would be known to the experimenter according to a randomly paired listing that was developed prior to the experimental session.

At the end of the testing session, participants received the dollar equivalent of their points in the form of a Best Buy gift voucher. The dollar-points equivalence table was constructed such that even if a subject earned no points during the game, they still received a minimum of \$15 across games. However, subjects did not know during the task that they would receive rewards even if their partners did not share points with them. Subjects received the dollar equivalent of points only at the end of the completed testing session. This type of research requires that compensation be tied to performance, as monetary motivation and incentive have substantial effects on the decision-making behavior within this task and must be maintained across participants (Camerer & Hogarth, 1999).

Coding justifications and attributions. After each decision to share or keep a preferred number of points, subjects were asked a series of qualitative questions to probe the social–cog-

nitive reasoning associated with each behavioral choice. First, the investor was asked to report the intention behind the initial offer to the Trustee (“Why did you make this choice?”). A careful content analysis of responses revealed five modal response types that could be reliably coded (a) fair intention (e.g., “It would be fair to give him half of my points. That way we start out with even footing—everyone wins” or “Because I thought it would be fair”), (b) intention to elicit cooperation (“Because if I’m generous and trust the trustee will probably be generous back” or “Because they might be generous enough to give the equal amount or more back to me”), (c) intention to mistrust (“Because I don’t know if I can trust my partner” or “I didn’t want to give away too many. Who knows if my partner will return any?”), (d) hostile intentions (“I gave him nothing so that I can have more points for myself” or “Because I want all the points for myself”), and (f) prosocial intentions (“Because I want to be kind” or “Because I’m feeling generous”). Responses were independently coded by coders who were provided with descriptions and examples of these categories for training purposes.

Second, upon receipt of the investor’s offer, the trustee was asked to evaluate the investor’s offer (“What do you think about the offer that has been made to you by the Investor?”). After a similar content analysis of responses, two modal categories of attributions emerged to categorize answers to this question: (a) fair offer (e.g., “That was a generous offer” or “He was generous so I will be”) and (b) unfair offer (e.g., “Greedy little miser” or “I don’t like him”).

Third, the trustee was asked to justify his return offer (“Why did you make this choice?”). The same five categories described for the investor justification question were used to categorize trustee responses.

Fourth, upon receipt of the return offer of the trustee, the investor was asked to evaluate the trustee’s return offer. Again, two modal categories of attributions emerged from the qualitative responses: (a) fair offer (e.g., “That was nice” or “That was kind of him”) and (b) unfair offer (e.g. “I think that he is a greedy dirt bag” or “I don’t like him”). The average interrater agreement across these four categories of responses was $\kappa = 0.82$ – 0.97 , indicating “near perfect interrater agreement” (Cohen, 1960).

Condition 2: “Known identity” trust game procedure. Prior to the start of the second experiment, children were told that they had been randomly assigned new partners for the second experiment. This time, photos of their partners appeared on the decision form. They were told that they would play the same game, and that they would remain in the role of the investor, but this time, they would know the identity of the person they were playing with.

The same procedure for decision making and eliciting justifications and attributions was followed, also for boys in the trustee role who were now aware of the identity of investors. The same four qualitative questions were asked after each decision and only one additional content category was needed to code the justifications given. Because the identity of partners

was known to boys in this condition, many boys justified their offers in both roles by referring to either the positive (e.g., "He's a cool guy") or negative (e.g., "He is a nasty person") attributes of their partners. These categories (positive attributes or negative attributes) were therefore added as categories for the known identity condition of the game.

Theory of mind

The CET (Baron Cohen et al., 2001) is an "off-line" theory of mind task and based on a deficit approach to theory of mind or social cognition (Fonagy & Sharp, 2008). It was adapted from the adult "Reading the Mind in the Eyes Test" developed by Baron-Cohen, Jolliffe, Mortimore, and Robertson (1997). The test includes 28 black and white photographs of the eye region and the participant is asked to pick one of four words that best describes what the person in the photo is thinking or feeling. Three of the four words are foil mental state terms, whereas the fourth is deemed "correct." The position of the four words is randomized for each item. Correct answers are scored as 1 and then summed to produce a total "correct" score. Other studies have used the CET in examining theory of mind and externalizing problems in children (Sharp, 2008; Sutton, Reeves, Keogh, et al., 2000), and in comparing theory of mind in children with Asperger symptoms and those without (Baron-Cohen et al., 2001). In a community sample of children aged 7 to 11, males scored at a mean of 17 ($SD = 3.81$), whereas females scored at a mean of 18 ($SD = 3.15$; Sharp, 2008). In contrast to the social-cognitive measure described above, the CET does not tap into on-line reporting of intentions of own behavior or the reading of intentions of partners' behavior.

Externalizing behavior problems

It is well known that different sources (e.g., parents, peer, and children themselves) often disagree on the presence and severity of problem behaviors (Verhulst & van der Ende, 1992). Low cross-informant correlations have often led researchers to cast doubt on one or both informants and have also been equated with unreliability (Gould, Bird, & Jaramillo, 1993). Still, it is important to keep in mind that different informants may validly contribute different information (Achenbach, McConaughy, & Howell, 1987; Verhulst & Van der Ende, 1992). Multiple informants are needed to obtain a comprehensive picture of an individual's functioning (Verhulst & van der Ende, 1992). To this end, we combined three measures of externalizing behavior to include youth self-report, parent report, and peer nominations.

Youth Self-Report (YSR) and parent report. The YSR (Achenbach & Rescorla, 2001) and Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) are well-established evidence-based assessment instruments (Holmbeck et al., 2008) that assess global and specific psychopathology among youth ages 6 to 18 years. It assesses behavioral and emotional symptoms during the past 6 months. The measure contains

112 problem items, each scored on a 3-point scale: 0 (*not true*), 1 (*somewhat or sometimes true*), to 2 (*very or often true*). The Total Problems Scale yields a *T* score of general psychiatric functioning and two broad subscales of externalizing behavior problems and internalizing behavior problems. Achenbach and Rescorla (2001) recommended using a *T* score cutoff at or above 65 to separate individuals at higher risk for psychopathology. Prior research indicates this threshold discriminates well between clinical and nonclinical populations (Achenbach & Rescorla, 2001). For the current study, we used *T* score of 65 on the externalizing subscale to group youth in above and below cutoff categories for externalizing behavior problems. In the current sample, 27% of boys were above cutoff on the YSR, and 11% were above cutoff on the CBCL. As shown in Table 1, descriptive statistics on these measures in our sample were comparable to Achenbach and Rescorla's (2001) nonreferred sample.

Peer nomination as relationally aggressive or prosocial. A peer-nomination instrument developed by Werner and Crick (1999) was used to assess relational aggression and prosocial behavior. The measure consists of 24 items. Seven of the items tap into a relational aggression subscale. This subscale has been found to be highly reliable, with a Cronbach α value of 0.087. Nine items tap into a prosocial behavior subscale and have also been shown to be highly reliable (Cronbach $\alpha = 0.91$; Werner & Crick, 1999).

As recommended by Werner and Crick (1999), participants were provided with a group membership roster to be used during self-administration of the peer-nomination instrument. For each of the behavioral items, participants were instructed to nominate up to five peers who best fit each description. The number of nominations each participant received from his or her peers was summed for each item and totaled for each subscale.

Next, we used the 50th percentile to identify boys above and below the median for relationally aggressive and prosocial nominations. As is often the case with sociometric studies (Cillessen & Bukowski, 2000) many boys were nominated as both prosocial and relationally aggressive, so many boys fell above the 50th percentile for both subscales. For the analyses we were interested in comparing trust and trustworthiness in boys who were perceived as "extreme" indicated by being nominated as "only prosocial" versus boys nominated in both categories. Similarly, we were interested in comparing boys who were "only relationally aggressive" to boys in both categories. Of the 79 boys in the Investor role, 14 were identified as "prosocial only" and 14 were identified as "relationally aggressive only." In the Trustee role, 24 boys were identified as "prosocial only" and 9 as "relationally aggressive only."

Combined measure of externalizing behavior problems. Because normality assumptions for externalizing behavior disorder variables were violated (possibly due to the fact that the sample was drawn from the community), all variables were submitted to a normalizing transformation. A boy was considered to meet criteria for externalizing behavior prob-

lems if he was above cutoff for parent report and self-report externalizing problems and described as “only relationally aggressive.” Using this variable, boys in the externalizing group showed significantly higher means for both self-reported ($t = -7.82$, $df = 156$, $p < .001$) and parent-reported ($t = -7.49$, $df = 156$, $p < .001$) externalizing behavior problems compared to boys without externalizing problems.

Data analytic strategy

The mean for investor offers was calculated by taking the average proportion (out of 20 points) that investors sent to the trustee. The mean for trustee return offers was calculated by averaging the ratios of trustee offers to the initial offer sent by the investor. For example, if the investor’s tripled offer was $3 \times 10 = 30$ and the trustee decided to send back 10 out of 30 points, the ratio or *relative return offer* was 0.33. These means served as indices for “trust” (mean investor offer) and “trustworthiness” (mean trustee relative return offer). Preliminary analysis was carried out on the distribution of scores and their associations with demographic variables. Both trust and trustworthiness estimates appeared to be normally distributed and did not correlate with parental education or age.

The first aim of the study was to examine whether externalizing problems (self-reported, parent reported, and peer ascertained) were reflected in trust and trustworthiness behaviors. We were particularly interested in the differential modulating impact of known identity (vs. anonymity) of trust game partners. To this end, we ran one-way repeated-measures analyses of variance (ANOVAs) with the investor and trustee offers, respectively, with game version (anonymous vs. known identity) as the repeated measures factor and group status for externalizing disorder as a between-subject factor.

Our second aim was to examine whether the relationship among trust, trustworthiness, and externalizing behavior problems could be linked to and accounted for by anomalies in social cognition. We had two indicators of social cognition. For the first, zero-order correlations (Pearson r) between trust and trustworthiness variables and the total score of the CET were calculated.

For the second (on-line social-cognitive reasoning during the task), we separately analyzed intentions underlying subjects’ own behavior (“Why did you make this choice?”) and assumed intentions of partners’ behavior (“What do you think about the offer that has been made to you by the Investor?”). We divided participant responses into predominantly “benevolent intentions” (fair intention, intention to elicit cooperation, prosocial intentions, and attributions of positive characteristics) and “hostile intentions” (hostile intentions, intentions to mistrust, and attributions of negative characteristics) on the basis of the frequency of attributions. We then contrasted the mean scores of these types of responses in terms of associated trust/trustworthiness scores using independent sample t tests. For assumed intentions of partners’ behavior, subjects’ evaluation of outcome was classified as either fair versus unfair and analyzed for its relation to trust/trustworthiness scores using independent sample t tests. Relations between fair versus unfair

evaluations of offers and group status in terms of externalizing behavior were examined using independent sample t tests.

Results

Comparison of trust (investor initial offer) across conditions taking group status (externalizing vs. no externalizing problems) into account

The means and standard deviations of a one-way repeated ANOVA are presented in Table 2. There was a significant effect for condition (anonymous vs. familiar), Wilks’ $\lambda < 0.001$, $F(1, 78) = 13.71$, $p < .001$, multivariate $\eta^2 = 0.15$ (large effect size; Cohen, 1988). The interaction between “experiment” and group status was not significant, Wilks’ $\lambda > .05$, $F(1, 78) = 1.02$, $p > .10$, multivariate $\eta^2 = 0.01$. Inspection of the means in Table 2 shows that for both groups the means increased with the known identity condition of the trust game compared to the anonymous condition.

Comparison of trustworthiness (reciprocity; trustee relative return offer) across conditions taking group status (externalizing vs. no externalizing problems) into account

The means and standard deviations for a repeated one-way ANOVA for trustworthiness ratings are presented in Table 2. There was no condition effect, Wilks’ $\lambda = 0.11$, $F(1, 78) = 2.54$, $p > .05$, multivariate $\eta^2 = 0.03$. However, the interaction between condition and group status was significant, Wilks’ $\lambda = 0.02$, $F(1, 78) = 5.61$, $p < .05$, multivariate $\eta^2 = 0.07$ (moderate effect size; Cohen, 1988). Inspection of the means shows that although means in trustworthiness increased for boys without externalizing problems when playing with a known partner, means actually decreased for boys with externalizing problems. This interaction effect is represented visually in Figure 1.

Relationship among social cognition, trust, and trustworthiness

Theory of mind. Pearson correlations between trust and trustworthiness variables and the total score of the CET were nonsignificant for investor offers on the anonymous ($r = -.13$, $p = .24$) and known identity ($r = -.08$, $p = .47$) conditions. Correlations were also nonsignificant for trustee relative return offers for the anonymous ($r = -.19$, $p = .09$) and known identity ($r = -.001$, $p = .99$) conditions. These results suggest no relationship between trust behavior and social cognition as indexed by an emotion recognition or theory of mind task such as the CET.

On-line social-cognitive reasoning to justify choices (intentions underlying own behavior). Table 3 summarizes the percentage of boys in each response category for the question “Why did you make this choice?” for the anonymous and the known identity conditions. For the anonymous game, most boys (47.4%) in the investor role decided to share points on

Table 2. Descriptive statistics

	Trust ^a			Trustworthiness ^b	
	Age (years)	Anonymous (n = 79)	Known Identity (n = 79)	Anonymous (n = 79)	Known Identity (n = 79)
Total sample (n = 158)					
Min	8	0	0	0	0
Max	18	20	20	1	1
Mean	12.84	7.19	10.66	0.22	0.27
SD	1.80	5.26	6.95	0.24	0.28
Externalizing (n = 44)					
Min	8	0	0	0	0
Max	18	10	20	0.47	0.67
Mean	12.77	6.32	8.53	0.17	0.15
SD	1.85	4.96	6.73	0.15	0.18
No externalizing (n = 144)					
Min	10	0	0	0	0
Max	16	20	20	1	1
Mean	13.02	7.43	11.33	0.25	0.31
SD	1.65	5.36	6.94	0.30	0.37

Note: The trustee relative return offers are calculated by dividing the trustee return offer into the investor investment.

^aInvestor initial offer (range = 0–20).

^bTrustee relative return offer (range = 0–1).

the basis of a fair intention. It is interesting that, once the identity of partners was known, the modal justification was based on attribution of positive characteristics to their partners (37.2%).

In the anonymous condition, most boys in the trustee role based their choices on benevolent intentions with 34.7% doing so to elicit cooperation and 30.6% doing so to communicate fair intentions. This was also the case for the known identity condition, although several boys in this condition also based their choices on the attribution of positive characteristics to their partners (16.9%).

We grouped “benevolent intentions” (fair intentions, intention to elicit cooperation, prosocial intentions, attribution of positive characteristics) and “hostile intentions” (hostile intent, intent to mistrust, attributions of negative characteristics) together to form two groups for further analyses. For the anonymous condition 71.8% of boys in the investor role had benevolent intentions versus 28.2% with hostile intentions, whereas 76.4% of boys in the trustee role had benevolent intentions versus 23.6% with hostile intentions. For the known identity condition, 79.5% of boys in the investor role had benevolent intentions versus 20.5% with hostile intentions. Although in the trustee role 67.6% had benevolent intentions versus 32.5% with hostile intentions, comparison between the anonymous and known identity proportions of benevolent versus hostile intentions was nonsignificant.

Table 4 summarizes the results of independent sample *t* tests to investigate the relation between social-cognitive style (benevolent versus hostile) and trust behavior (investor offers; trustee relative return offers).

For both roles and for both games there were highly significant differences between boys with “benevolent” versus “hostile” intentions in game playing behavior. Boys with hos-

tile intentions were significantly less trusting ($M = 4.27$, $SD = 4.82$) than boys with benevolent intentions ($M = 8.45$, $SD = 4.96$) during the anonymous game. They were also less trustworthy ($M = 0.02$, $SD = 0.08$) compared to boys with benevolent intentions ($M = 0.26$, $SD = 0.23$).

These results mirrored results in the known identity condition of the game. Boys with hostile intentions were significantly less trusting ($M = 4.60$, $SD = 7.10$) than boys with benevolent intentions ($M = 12.27$, $SD = 6.15$). They were also less trustworthy ($M = 0.01$, $SD = 0.03$) compared to boys with benevolent intentions ($M = 0.38$, $SD = 0.24$).

Boys with hostile intentions were also more likely to have higher self-reported externalizing behavior than boys with benevolent intentions, although these mean differences were significant only for Trustees in the known identity condition ($M = 58.72$ vs. $M = 53.08$, $t = 0.12$, $df = 75$; $p = .02$).

On-line social-cognitive evaluation of outcomes. As shown in Table 3, by and large, evaluations of investor offers by trustee partners were positive, with more than two-thirds judging offers as fair for both games. Investor evaluations of trustee relative return offers were less favorable, with around half of boys judging these as fair.

We investigated whether an “unfair” judgment of Investor’s initial offers was related to trustee relative return offers. The results in Table 4 show that the relative return offers from trustees were significantly lower for boys who evaluated offers from investors as unfair ($M = 12$, $SD = 0.21$) compared to boys who viewed offers as fair ($M = 0.27$, $SD = 0.21$) in the anonymous game. Similar results were found for the known identity condition for boys who viewed offers as unfair ($M = 0.17$, $SD = 0.36$) versus boys with who viewed of-

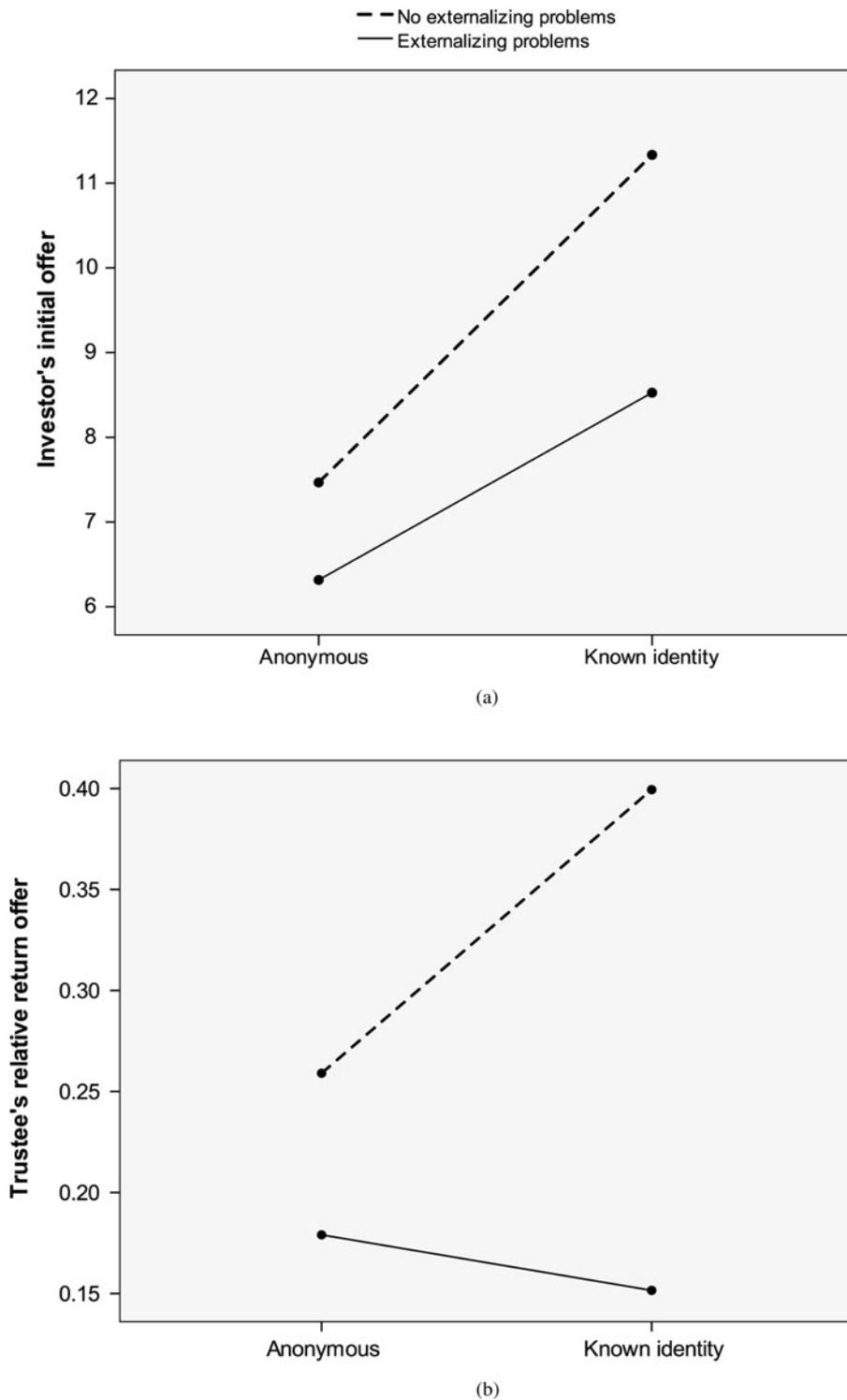


Figure 1. The interaction between experimental condition (anonymous vs. known identity) and group status (externalizing vs. nonexternalizing). (a) Trust and (b) trustworthiness.

fers as fair ($M = 0.34$, $SD = 0.23$). When the offer was judged as fair, trustees were significantly more likely to make higher relative return offers compared to when initial offers were judged as unfair. Judgments of offers as unfair were, however, not related to externalizing problems.

Discussion

Over the last three decades the role of social cognition in the development and maintenance of externalizing behavior problems has been a topic of much interest. Experimental para-

Table 3. Social-cognitive reasoning associated with behavioral choices and outcomes

	Anonymous		Known Identity	
	Category	%	Category	%
On-Line Social-Cognitive Reasoning to Justify Choices ^a				
The investor: Why did you make this choice?	Benevolent intentions		Benevolent intentions	
	Fair intention	47.4	Fair intention	17.9
	Intention to elicit cooperation	19.2	Intention to elicit cooperation	14.1
	Prosocial intentions	5.1	Prosocial intentions	10.3
			Positive characteristics	37.2
	Hostile intentions		Hostile intentions	
		Intention to mistrust	1.3	
		Malevolent intention	12.8	
		Negative characteristics	6.4	
The trustee: Why did you make this choice?	Benevolent intentions		Benevolent intentions	
	Fair intention	30.6	Fair intention	24.7
	Intention to elicit cooperation	34.7	Intention to elicit cooperation	20.8
	Prosocial intentions	11.1	Prosocial intentions	5.2
			Positive characteristics	16.9
	Hostile intention		Hostile intentions	
		Intention to mistrust	14.3	
		Malevolent intention	13	
		Negative characteristics	5.2	
On-Line Social-Cognitive Evaluation of Outcomes ^b				
The trustee: What do you think about the offer that has been made to you by the investor?	Fair offer	65.6	Fair offer	74.6
	Unfair offer	34.4	Unfair offer	25.4
The investor: What do you think about the return offer that has been made to you by the trustee?	Fair offer	50	Fair offer	53
	Unfair offer	50	Unfair offer	46

^aSubject's own intentions.

^bEvaluation of other player's intentions.

digms based on game theory principles offer a novel approach to operationalizing social-cognition in real-life, real-time, controlled settings. One such paradigm, the trust task, complements psychodynamic (Erikson, 1950) and social learning conceptualizations (Rotter, 1967, 1971) of trust by parametrically encoding trust and trustworthiness as the amount of money designated to partners during an economic exchange game (Montague, King-Casas, & Cohen, 2006). With the exception of a few studies (Harbaugh, Krause, Liday, & Vesterlund, 2003; Sutter & Kocher, 2007; van den Bos et al., in press), the trust task has not been used in youth samples. Further, bar one study (McClure et al., 2007), economic exchange games have not been used in relation to psychopathology in youth.

In the current study, we used the trust task to investigate anomalous trust and trustworthiness in boys with and without externalizing behavior problems. Several authors have argued that trust and trustworthiness are clinically important factors in psychopathy and the development of criminal behavior (Bernath & Feshbach, 1995), difficulties in achieving satisfaction in work and interpersonal relationships (Mowrer & Ullman, 1945), delinquency and social irresponsibility (Mischel, 1961), as well as cheating and yielding to temptations (Mischel & Gilligan, 1964). Bernath and Feshbach (1995) ar-

gued that children who fail to trust others may be at risk of developing an antagonistic defensive posture and may view social relations as "get them before they get you," thus priming irresponsible, potentially criminal and delinquent behavior. Although early work has provided some evidence for the relationship between abnormal *expectations* related to trust and externalizing problems (e.g., Wentzel, 1991; Wright & Kirimani, 1977), direct evidence in relation to anomalous observed social behavior is lacking.

This is the first study to empirically link trust and trustworthiness to externalizing behavior problems indexed by parent report, self-report, and peer report. It is also the first study (in adults or children) to link on-line social-cognitive reasoning and concurrent real-life trust behavior with externalizing behavior difficulties. Results supported the conclusion that boys with externalizing behavior problems are generally less trustworthy, but not necessarily less trusting than boys without externalizing problems. More specifically, our results showed that knowing the identity of one's partner increased trust in both groups of children. This was also true for trustworthiness in nonexternalizing boys. However, in boys with externalizing problems trustworthiness was reduced in the known identity condition. Moreover, reduced trust and trustworthiness

Table 4. Social cognition and game playing behavior

Game	Behavior	Social-Cognitive Style	Mean (SD)	<i>t</i>	<i>df</i>	<i>p</i>
On-Line Social-Cognitive Reasoning to Justify Choices ^a						
Anonymous	The investor initial offer	Benevolent intentions	8.45 (4.96)	3.37	76	.001
		Hostile intentions	4.27 (4.82)			
Known identity	The investor initial offer	Benevolent intentions	12.27 (6.15)	4.20	75	<.001
		Hostile intentions	4.60 (7.10)			
Anonymous	The trustee relative return offer	Benevolent intentions	0.26 (0.23)	4.04	70	<.001
		Hostile intentions	0.02 (0.08)			
Known identity	The trustee relative return offer	Benevolent intentions	0.38 (0.24)	7.35	75	<.001
		Hostile intentions	0.01 (0.03)			
On-Line Social-Cognitive Evaluation of Outcomes ^b						
Anonymous	The trustee relative return offer	Fair	0.27 (0.21)	2.78	62	.007
		Unfair	0.12 (0.21)			
Known identity	The trustee relative return offer	Fair	0.34 (0.23)	2.17	61	.03
		Unfair	0.17 (0.36)			

Note: Mean (SD) refers to the means of investor (trust) and trustee offers (trustworthiness) divided according to their social-cognitive reasoning style (benevolent vs. hostile intentions; fair vs. unfair evaluation of outcomes).

^aSubject's own intentions.

^bEvaluation of other player's intentions.

were associated with hostile intentions underlying subjects' own behavior as well as judgment of partners' behavior as unfair, which in turn related to externalizing behavior problems. Finally, although untrusting and untrustworthy behaviors were associated with hostile intentions and judgments of partners' behaviors as unfair, such behaviors were not associated with a general theory of mind deficit based on judgments of facial expressions of internal states as measured by the CET.

By demonstrating the link between on-line social-cognitive reasoning and concurrent real-life trust game behavior, we provide support for models of trust game behavior that fall within the class of intention-based accounts for adult (Dufwenberg & Gneezy, 2000; Falk & Fischbacher, 2006; McCabe et al., 2000) and youth samples (Van den Bos et al., in press). These models emphasize players reading each other's intentions or motives (and not merely their actions). In contrast, outcome-based models focus simply on behavior and have suggested that a certain proportion of the population is altruistic or spiteful or have certain thresholds of inequity aversion (Bolton & Ockenfels; Fehr & Schmidt, 1999). The fact that boys with externalizing behavior problems showed problems in trustworthiness, but not trust, mitigates against a simple outcome-based interpretation of our results. After all, if boys with externalizing disorder were simply less altruistic or spiteful in their behavior, such behavior should have been evident in their Investor initial offers (trust). Instead, they show anomalies only in trustworthiness or *reciprocity*. Further, these anomalies appear to be associated with on-line hostile intentions.

Our results build on, and are explained by, findings from previous studies of social cognition and externalizing behavior problems, specifically findings demonstrating hostile attribution biases in boys with conduct problems (for a review,

see Mize & Pettit, 2008). The current study expands this research by showing that on-line hostile intentions underlying subjects' own behavior are associated with real-time, real-life behavior during a social exchange game; and that boys with externalizing behavior problems differ from normal controls both in their social-cognitive reasoning and associated game behavior. Our data cannot speak to why boys with externalizing behavior problems are less trustworthy or more hostile in their intentions. It is, however, worth speculating that the development of anomalous trust behavior associated with hostile social-cognitive reasoning may have its foundation in insecure attachment (Fonagy, Gergely, & Target, 2007; Mize & Pettit, 2008). Insecure attachment has been shown to relate to externalizing behavior problems (Green, Stanley, & Peters, 2007; Greenberg et al., 2001; Warren, Emde, & Sroufe, 2000; Warren, Oppenheim, & Emde, 1996). Attachment in turn plays a role in the development of various social-cognitive capacities, including perspective taking, theory of mind, and mentalizing (Sharp & Fonagy, 2008), potentially mediating the development of trust (Fonagy & Target, 1997). To test this hypothesis, future research should focus on including simultaneous measures of attachment, social-cognitive reasoning, and trust in samples of youth with externalizing behavior problems.

To understand why the known identity condition was associated with higher trustworthiness in normal controls, but lower trustworthiness in externalizing boys, we turn to literature investigating the effect of known identity on trust game behavior. Studies in adults have demonstrated that prior social and moral information about partners in the trust game modulate responses during economic exchange games. For instance, Delgado, Frank, and Phelps (2005) showed that prior social and moral information was associated with level of trust in partners, as well as actual sharing behavior. It has also been

shown that knowing the identity of partners or opponents attenuates aggressive acts and promotes trust and trustworthiness during the trust game (Camerer, 2003), because knowing the identity of partners produces a greater empathic response accompanied by a greater willingness to make personal sacrifices to benefit another (Jenni & Loewenstein, 1997). It is possible that this effect did not occur in boys with externalizing behavior problems due to reduced affective empathy demonstrated for this group of children (Lovett & Sheffield, 2007). Future studies in which anonymous and known identity versions of the trust task are administered alongside measures of empathy may further support this hypothesis.

The current study has several limitations. First, the sample was community based, which afforded us the opportunity to exploit real-life established relationships between boys. Despite our efforts to demonstrate comparability between the current sample and other community samples, it is possible that boys who self-select into the Scouts might differ in some ways from the general population. Therefore, the results should be replicated in other community samples. Moreover, this study aimed to elucidate potential disease mechanisms for externalizing behavior problems and should therefore also be replicated in a sample of clinically referred youth. Second, although 79 pairs afforded us enough power to detect differences, a larger sample size (combined with clinical characteristics) will result in larger effect sizes, thereby more strongly establishing the role of trust and trustworthiness in externalizing behavior problems. Third, we could not control for the past history of the relationship quality between partners, and we collected no information about this. We interpret our behavioral results as an indication of a bias of interpersonal behavior in externalizing individuals based on social cognition. The source of this bias may be based, in part, on past history of a relationship, for instance, friendship quality.

References

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implication of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.
- Ashraf, N., Bohnet, I., & Plankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, *9*, 193–208.
- Axelrod, R. (1986). An evolutionary approach to norms. *American Political Science Review*, *80*, 1095–1011.
- Baron Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger Syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*, 813–822.
- Baron Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the mind in the eyes,” test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*, 241–251.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*, 5–17.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social-history. *Games and Economic Behavior*, *10*, 122–142.
- Bernath, M. S., & Feshbach, N. D. (1995). Children’s trust: Theory, assessment, development and research directions. *Applied and Preventive Psychology*, *4*, 1–19.
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, *90*, 166–193.
- Boy Scouts of America. (2010). *Scoutreach*. Retrieved from <http://www.scouting.org/About/AnnualReports/PreviousYears/2002/scoutreach.aspx>
- Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.
- Camerer, C., & Hogarth, R. M. (1999). The effects of financial incentives in economics experiments: A review and capital–labor–production framework. *Journal of Risk and Uncertainty*, *18*, 7–42.
- Camerer, C. F., & Weigelt, K. (1988). Experimental tests of the sequential equilibrium reputation model. *Econometrica*, *56*, 1–36.
- Cillessen, A. H. N., & Bukowski, W. M. (2000). *Recent advances in the measurement of acceptance and rejection in the peer system*. San Francisco, CA: Jossey-Bass.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

The anomalous behavior of externalizing children may be based on the increased likelihood of actual negative relationship histories of externalizing children relative to normally developing children. Of course, the social–cognitive biases pointed to in this study may have played a role in creating such negative histories. Future studies will need to control for this important confounding effect. Fourth and finally, the two games were not counterbalanced. Thus, there is no way of knowing whether trusting behavior naturally increased as subjects gained more experience in playing the game. A replication study should therefore include counterbalancing.

Setting these limitations aside, this study identified the potential of economic exchange tasks in studying SIP biases in relation to clinical problems. We have provided concurrent validity for the behavioral measure by using more traditional verbal self-report measures of social–cognitive bias. The key advantage of methods such as the trust task is that it parameterizes aspects of abnormal interpersonal behavior that are central to many disorders in developmental psychopathology. The application of this approach opens the door to behavioral measures of cognitive biases currently inadequately measured by paper-and-pencil tests of limited ecological validity. It also has potential for the computational modeling of the anomalies associated with these disorders with knock-on advantages for interfacing with brain-based measurements (Montague et al., 2006). Further, if our initial demonstration of anomalies is replicated in future investigations, studies will be able to examine precursors of trust and trustworthiness problems by, for instance, addressing issues of causation in relation to past interpersonal experiences. Moreover, follow-up studies may focus on the malleability of these anomalies by psychosocial or pharmacological interventions, for example, addressing the question of whether parenting training approaches result in changes in trustworthiness.

- Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, 8, 1611–1618.
- Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development*, 51, 162–170.
- Dodge, K. A. (1993). Social-cognitive mechanisms in the development of conduct disorder and depression. *Annual Review of Psychology*, 44, 559–584.
- Dodge, K. A., Lansford, J. E., Burks, V. S., Bates, J. E., Pettit, G. S., Fontaine, R., et al. (2003). Peer rejection and social information-processing factors in the development of aggressive behavior problems in children. *Child Development*, 74, 374–393.
- Dodge, K. A., & Somberg, D. R. (1987). Hostile attributional biases among aggressive boys are exacerbated under conditions of threats to the self. *Child Development*, 58, 213–224.
- Dufwenberg, M., & Gneezy, U. (2000). Measuring beliefs in an experimental wallet game. *Games and Economic Behavior*, 30, 163–182.
- Erikson, E. (1950). *Childhood in society*. New York: Norton.
- Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54, 293–315.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114, 817–868.
- Fonagy, P., Gergely, G., & Target, M. (2007). The parent-infant dyad and the construction of the subjective. *Journal of Child Psychology and Psychiatry*, 48, 288–328.
- Fonagy, P., & Sharp, C. (2008). Treatment outcome of childhood disorders: The perspective of social cognition. In C. Sharp, P. Fonagy & I. M. Goodyer (Eds.), *Social cognition and developmental psychopathology*. Oxford: Oxford University Press.
- Fonagy, P., & Target, M. (1997). Attachment and reflective function: Their role in self-organization. *Development and Psychopathology*, 9, 697–700.
- Fonagy, P., Twemlow, S. W., Vernberg, E. M., Nelson, J. M., Dill, E. J., Little, T. D., et al. (2009). A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. *Journal of Child Psychology and Psychiatry*, 50, 607–616.
- Gould, M. S., Bird, H., & Jaramillo, B. S. (1993). Correspondence between statistically derived behavior problem syndromes and child psychiatric diagnoses in a community sample. *Journal of Abnormal Child Psychology*, 21, 287–313.
- Green, J., Stanley, C., & Peters, S. (2007). Disorganized attachment representation and atypical parenting in young school age children with externalizing disorder. *Attachment and Human Development*, 9, 207–222.
- Greenberg, M. T., Speltz, M. L., DeKlyen, M., & Endriga, M. C. (1991). Attachment security in pre-schoolers with and without externalizing behavior problems: A replication. *Development and Psychopathology*, 3, 413–430.
- Greenberg, M. T., Speltz, M. L., DeKlyen, M., & Jones, K. (2001). Correlates of clinic referral for early conduct problems: Variable- and person-oriented approaches. *Development and Psychopathology*, 13, 255–276.
- Harbaugh, W. T., Krause, K., Liday, S. G., & Vesterlund, L. (2003). Trust in children. In E. Ostrom & J. Walker (Eds.), *Trust, reciprocity and gains from association: Interdisciplinary lessons from experimental research*. New York: Russell Sage Foundation.
- Holmbeck, G. N., Thill, A. W., Bachanas, P., Garber, J., Miller, K. B., Abad, M., et al. (2008). Evidence-based assessment in pediatric psychology: Measures of psychosocial adjustment and psychopathology. *Journal of Pediatric Psychology*, 33, 958–980, Discussion 981–952.
- Hughes, C., & Ensor, R. (2008). Does executive function matter for preschoolers' problem behaviors? *Journal of Abnormal Child Psychology*, 36, 1–14.
- Jenni, K. E., & Loewenstein, G. (1997). Explaining the "identifiable victim effect." *Journal of Risk and Uncertainty*, 14, 235–257.
- Lovett, B. J., & Sheffield, R. A. (2007). Affective empathy deficits in aggressive children and adolescents: A critical review. *Clinical Psychology Review*, 27, 1–13.
- Malhotra, D. (2004). Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties. *Organizational Behavior and Human Decision Processes*, 94, 61–73.
- McCabe, K. A., Smith, V. L., & LePore, M. (2000). Intentionality detection and "mindreading": Why does game form matter? *Proceedings of the National Academy of Sciences of the United States of America*, 97, 4404–4409.
- McClure, E. B., Parrish, J. M., Nelson, E. E., Easter, J., Thorne, J. F., Rilling, J. K., et al. (2007). Responses to conflict and cooperation in adolescents with anxiety and mood disorders. *Journal of Abnormal Child Psychology*, 35, 567–577.
- Mischel, W. (1961). Preference for delayed reinforcement and social responsibility. *Journal of Abnormal and Social Psychology*, 62, 1–7.
- Mischel, W., & Gilligan, C. (1964). Delay of gratification, motivation for the prohibited gratification, and resistance to temptation. *Journal of Abnormal and Social Psychology*, 69, 411–417.
- Mize, J., & Ladd, G. W. (1988). Predicting preschoolers' peer behavior and status from their interpersonal strategies: A comparison of verbal and enactive responses to hypothetical social dilemmas. *Developmental Psychology*, 24, 782–788.
- Mize, J., & Pettit, G. S. (2008). Social information processing and the development of conduct problems in children and adolescents: Looking beneath the surface. In C. Sharp, P. Fonagy & I. M. Goodyer (Eds.), *Social cognition and developmental psychopathology* (pp. 141–174). Oxford: Oxford University Press.
- Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience*, 29, 417–448.
- Mowrer, O. H., & Ullman, A. D. (1945). Time as a determinant in integrative learning. *Psychology Review*, 52, 61–90.
- Orobio de Castro, B., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouer, H. J. (2002). Hostile attribution of intent and aggressive behavior: A meta-analysis. *Child Development*, 73, 916–934.
- Pillutla, M. M., Malhotra, D., & Murnighan, J. K. (2003). Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology*, 39, 448–455.
- Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, 35, 651–665.
- Rotter, J. B. (1971). Generalized Expectancies for Interpersonal Trust. *American Psychologist*, 443–452.
- Sharp, C. (2008). Theory of mind and conduct problems in children: Deficits in reading the "emotions of the eyes." *Cognition and Emotion*, 22, 1149–1158.
- Sharp, C., Croudace, T. J., Goodyer, I. M. (2007). Biased mentalizing in children aged seven to 11: Latent class confirmation of response styles to social scenarios and associations with psychopathology. *Social Development*, 16, 181–202.
- Sharp, C., & Fonagy, P. (2008). Social cognition and attachment-related disorders. In C. Sharp, P. Fonagy, & I. Goodyer (Eds.), *Social cognition and developmental psychopathology* (pp. 269–302). Oxford: Oxford University Press.
- Sharp, C., Fonagy, P., & Goodyer, I. M. (2008). Introduction: Social cognition and developmental psychopathology. In C. Sharp, P. Fonagy, & I. M. Goodyer (Eds.), *Social cognition and developmental psychopathology*. Oxford: Oxford University Press.
- Singer, T., & Fehr, E. (2005). The neuroeconomics of mindreading and empathy. *AEA Papers and Proceedings*, 95, 340–345.
- Sutter, M., & Kocher, M. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior*, 59, 364–382.
- Sutton, J., Reeves, M., & Keogh, E. (2000). Disruptive behaviour, avoidance of responsibility and theory of mind. *British Journal of Developmental Psychology*, 18, 1–11.
- Van den Bos, W., Westenberg, M., Van Dijk, E., & Crone, E. A. (in press). Development of trust and reciprocity in adolescence. *Cognitive Development*.
- Verhulst, F. C., & Van der Ende, J. (1992). Agreement between parents' reports and adolescents' self-reports of problem behavior. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 33, 1011–1023.
- Vitaro, F., Tremblay, R. E., & Bykowski, W. M. (2001). Friends, freinships and conduct disorders. In J. Hill & B. Maughan (Eds.), *Conduct disorders in childhood and adolescence* (pp. 346–376). Cambridge: Cambridge University Press.
- Warren, S. L., Emde, R. N., & Sroufe, L. A. (2000). Internal representations: Predicting anxiety from children's play narratives. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 100–107.
- Warren, S. L., Oppenheim, D., & Emde, R. N. (1996). Can emotions and themes in children's play predict behavior problems? *Journal of the American Academy of Child & Adolescent Psychiatry*, 35, 1331–1337.
- Wentzel, K. R. (1991). Relations between social competence and academic achievement in early adolescence. *Child Development*, 62, 1066–1078.
- Werner, N. E., & Crick, N. R. (1999). Relational aggression and social-psychological adjustment in a college sample. *Journal of Abnormal Psychology*, 108, 615–623.
- Wright, T. L., & Kirmani, A. (1977). Interpersonal trust, trustworthiness and shoplifting in high school. *Psychological Reports*, 41, 1165–1166.