

Learning whom to trust in repeated social interactions: A developmental perspective

Group Processes & Intergroup Relations

1–14

© The Author(s) 2011

Reprints and permission: sagepub.

co.uk/journalsPermissions.nav

DOI: 10.1177/1368430211418698

gpir.sagepub.com



Wouter van den Bos,^{1,2,3} Eric van Dijk^{1,2}
and Eveline A. Crone^{1,2,4}

Abstract

How do people learn to trust or distrust others? In a repeated trust game setting, we investigated the development of trust within repeated interactions. We assessed the development of relation-specific trust across different age groups, ranging from late childhood to young adulthood. The results demonstrated that within relations the older participants showed lower levels of negative reciprocity. Additionally, with increasing age both the anger towards, and punishment of, noncooperative players decreased. Further analyses showed that the differential willingness to punish violations of trust was mediated by feelings of anger. Overall, the data provide initial evidence for the role of emotion regulation in the development of interpersonal trust.

Keywords

trust, development, anger, punishment, decision making

Trust plays an important role in almost all types of social and organizational interactions (Rotter, 1967). In our daily lives trust is important in our relationships with family and friends, and for groups and organizations, it reduces transaction costs (Barney & Hansen, 1994) and facilitates alliances (Gulati, 1995). Trust is important at all levels of society; it is often considered the “glue” that holds society together (Fukuyama, 1995; Sullivan & Transue, 1999; Zak & Knack, 2001), and trust is recognized to be of great importance for the development of social functioning throughout life, promoting moral behavior (Wright & Kirmani, 1977) and academic achievement (Imber, 1973; Wentzel, 1991).

Although it is clear that trust has many benefits, it “is an equal failing to trust everybody, and to trust nobody.” This raises the question how we learn to trust *or* distrust persons and institutions? Several factors have been identified that trust (or

¹Leiden University

²Leiden Institute for Brain and Cognition

³Stanford University

⁴University of Amsterdam

Corresponding author:

Wouter van den Bos, Department of Psychology, Stanford University, Jordan Hall, 450 Serra Mall, CA 94305, Stanford, USA

Email: wvdbos@stanford.edu

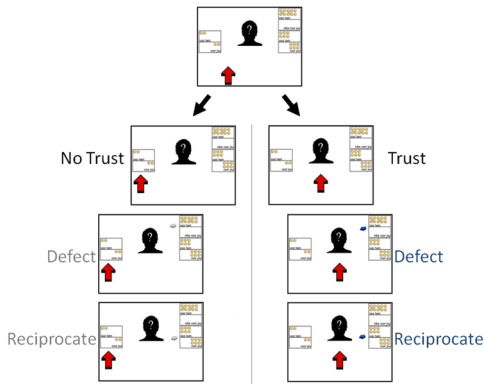


Figure 1. The sequence of visual displays that represent the different stages of the Simultaneous Trust Game (DTG). Each round starts with the identification of the other player. Next, the participant can decide to either trust or not to trust the other player. Regardless of the decision of the participant the choice by the other player to reciprocate or defect will be revealed. If the player decided to trust the choice by the other player is indicated with blue arrows and represent the real outcome of the game, whereas if the participant decided not to trust these arrows are displayed in grey and represent the counterfactual outcome of the game. Note that in the real experiment photographs of other players were shown, instead of the silhouette with a question mark.

distrust) is based on, including: an individual's general propensity to trust (Mayer, Davis, & Schoorman, 1995; Rotter, 1967, 1971), relationship history (Boon & Holmes, 1991; Boyle & Bonacich, 1970) and/or relationship-specific emotions (Dunn & Schweitzer, 2005; Huang & Murnighan, 2010; McAllister, 1995). It is hypothesized that individual characteristics, such as personality and developmental history, influence the initial propensity to trust (Mayer et al., 1995; Rotter, 1971; Yamagishi, Cook, & Watabe, 1998), whereas relation-specific trust, and related emotions (Dunn & Schweitzer, 2005), are based on a series of interactions with a person or institution.

To study the development of trust, social psychology has made extensive use of the trust game (TG). In the TG there are two players who can share a certain amount of money. The first player (trustor) has the possibility to divide a sum

of money equally or to give it all to the second player (trustee). If the first player decides to share the money, both players get their equal share and the game ends. However, if the first player gives all the money to the second player the total amount of money is tripled. Next, the second player has the possibility to reciprocate trust and share the increased amount of money with first player, or to exploit trust by keeping all the money (see Figure 1). It is clear that in the TG, Player 1 faces the challenging question of whether or not to trust Player 2: Will he/she reciprocate an act of trust? In more general terms, such an act of trust can be defined as the "willingness to make oneself vulnerable to others' actions based on a certain expectation of positive reciprocity" (Colman, 2003).

Interestingly, research shows that the initial trust in others is often high (Berg, Dickhaut, & McCabe, 1995; Dufwenberg & Gneezy, 2000; McCabe, Houser, Ryan, Smith, & Trouard, 2001). That is, in one-trial settings people often show high levels of trust, and in repeated settings they often show trust on the first encounter. Furthermore, research in which adults played multiple trust games has revealed that trust in others generally increases after positive trust experiences, and decreases after experienced violations of trust (e.g., Camerer & Ho, 1999; Erev & Roth, 1998). Models of trust learning suggest that relatively simple reinforcement-learning mechanisms update the expected outcome of future interactions. This hypothesis is supported by recent neuroimaging studies using multiple-round trust games (e.g., Delgado, Frank, & Phelps, 2005; King-Casas et al., 2005). Interestingly, these studies also suggested that besides updating the expected value, the affective responses to reciprocated or violated trust play an important role in the subsequent decisions to trust (King-Casas et al., 2008).

Recent research within the field of developmental psychology (Sutter & Kocher, 2006; Harbaugh, Krause, Liday, & Vesterlund, 2002; van den Bos, Westenberg, van Dijk, & Crone, 2010) has shown that there are still important age-related changes in general trust. These

studies indicate that adults—that is, the typical participants in social psychology studies—have higher levels of general trust compared to children. Whereas these findings address an important aspect of the development of trust, they do not inform us on age-related changes in relation-specific trust.

Learning to trust and distrust

So how does trust develop? From the above, it is clear that to answer this question, we should distinguish between the general trust and relation-specific trust. It is also clear that social psychology and developmental psychology have each addressed different parts, but currently, the literatures more or less stand alone. To provide a more comprehensive picture of the development of trust, we therefore set out to integrate insights from both fields. For this purpose we use a repeated trust game paradigm in which participants from different ages (children, adolescents, and adults) interact with the same players for several rounds (King-Casas et al., 2005).

Because, as Rotenberg (1980) emphasized, it is equally important to learn whom *not* to trust as to learn whom to trust, the participants in the current experiment interacted with three different preprogrammed personalities that displayed different levels of trustworthiness (low, medium, and high). During the repeated interactions the participants were playing the role of the trustor, thus each round they had to decide whether or not to trust the other. Following the number of trust decisions of the participants, we were able to study how the level of trust for each player changed based on the outcome of a series of social interactions.

Based on previous developmental studies with one-shot games we expect that with age participants will show higher levels of general trust and thus will be more prone to start the interaction with a trust move (Berg et al., 1995; Sutter & Kocher, 2008; van den Bos et al., 2010). Subsequently, based on the outcomes of the social interaction with the three different players we expect that participants will learn how

trustworthy each of the players is, and will act accordingly. We propose that although children display low levels of general trust on the first encounter, they are able to learn to trust *and* distrust their interaction partners based on a series of interactions. However, we expect that children will show a different learning strategy than adults, particularly being more sensitive to violations of trust.

Research from the domain of developmental psychology, using questionnaires and self-reports, suggests that trust relationships with peers already exist at a young age, but are initially very fragile and become more stable over the years. These studies on social relationships show that between ages 8 and 11, children's estimation of trustworthiness are sensitive to recent violations of trust, whereas only at a later age trust is increasingly based on consistent patterns of behavior over time (Rotenberg & Pilipenko, 1983–1984). At the latest stage of development, starting around early adolescence (11–13 years of age) and lasting until late adolescence, friendships become increasingly stable and resistant to violations of trust (Kahn & Turiel, 1988). In addition, it is well known that children are less capable to regulate their emotions in social situations than adults (Eisenberg, 2000), and emotion regulation is thought to develop until at least late adolescence/young adulthood (Casey, Jones, & Hare, 2008). Therefore, we expect that children, compared to adults, are particularly sensitive to trust violations. As a result, children are expected to more often decide not to trust in the round following a trust violation, whereas adults were expected to be more forgiving based on the history of previous interactions.

Trustworthiness, anger, and punishment

To further investigate the role of emotions and emotion regulation in the ontogeny of relation-specific trust, we explored the relation between emotional reactions to violated trust, and subsequent (costly) punishment of the violator. Unreciprocated trust and noncooperative

behavior in general are known to cause personal distress and, in particular, anger towards the non-cooperator (Pillutla & Murnighan, 1996; Seip, van Dijk & Rotteveel, 2009; Stouten, De Cremer, & van Dijk, 2009). In addition, it is often assumed that the anger towards uncooperative norm violators, in this case of the norm of reciprocity (Gouldner, 1960), may motivate people to punish the perpetrator (Pillutla & Murnighan, 1996), even when this punishment is costly (Fehr, 2002; Fehr & Fischbacher, 2003). Although there is some evidence for a causal relation between anger and punishment, this is not yet well established (see Seip et al., 2009). Furthermore, to our knowledge there are currently no studies that have investigated the relation between negative affect and costly punishment in developmental populations.

First of all, because children are less capable to regulate their emotions in social situations than adults (Eisenberg, 2000), we expect that the anger evoked by untrustworthy behavior will be higher for children than for adults. The increased anger could in turn lead to an increase in the level of punishment. This hypothesis is supported by studies that show that reduced self-regulation is strongly related to increased levels of reactive aggression (Connor, Steingard, Cunningham, Anderson, & Melloni, 2004; Winstok, 2009). Reactive aggression is a particular form of aggressive behavior that is evoked by perceived threat or provocation (Dodge & Coie, 1987), in this experiment the violation of trust.

Second, circumstantial evidence for our hypothesis that children will punish noncooperators more than adults comes from developmental studies with the Ultimatum Game. In these studies participants are offered a split of a certain amount of money between themselves and another player. The results of these studies show that children reject unfair offers (unequal splits in advantage of the other player) more often than adults do (Murnighan & Saxon, 1998; Güroğlu, van den Bos, & Crone, 2009; Sutter, 2007). Such rejections have been interpreted as means to punish, as they directly reduce the outcomes of the proposer.

In sum, there is some evidence for higher levels of anger and punishment in children

compared to adults. However, no previous study investigated the relation between these two concepts in developmental populations. To investigate the relation between negative affect and costly punishment, we will measure the participants' feelings of anger towards the other players and their use of (costly) punishment (cf. Fehr, 2002) after they have finished the trust game.

Method

Sample

Our sample included 60 participants (30 male, 30 female) divided over three age groups; late childhood (M age = 11.33, SD = 0.48, 9 male, 9 female), midadolescence (M age = 16.24, SD = 0.91, 13 male, 8 female) and young adulthood (M age = 21.06, SD = 2.27, 8 male, 13 female). Chi-square analyses indicated that gender distributions did not differ significantly between age groups, $\chi^2(3) = 5.69, p = .078^1$. Children and adolescents were recruited by contacting local schools. Child and adolescent participants were selected with the help of their teachers (children with learning or psychiatric disorders were excluded); informed consent was obtained from a primary caregiver. Adults were recruited at the university.

Simultaneous trust game

To study how participants learn whom to trust or distrust in a trust game setting we employed the Simultaneous Trust Game (STG) with repeated interactions. In the STG (Figure 1) the participants played multiple trust games in which both players simultaneously had to make their decision. Participants played the STG with three different players in a single session. At the start of each round the screen displayed the photograph of the other player, who was always matched for age and gender. Next, the participant saw the complete decision tree and had to choose from two options: to trust or not to trust. If the participant decided not to trust, the coins were divided evenly, one euro each, between the players. If the participants decided to trust the other player the

total money in the game was tripled in value (new total €3.00). When the other player had decided to reciprocate, the €3.00 were again divided evenly, resulting in €1.50 for each player. However, if the other player decided to defect she would take all the €3.00 and leave the participant with nothing. The payoff structure of the game was the same for every round (see Figure 1).

In the STG both players independently made their decision before the decision of the other is revealed, and in the end both decisions were always revealed. Thus before the decision of the participant to trust or not is revealed, the other player already had to decide if she would share or take all the money *if* she was trusted by the participant. Because the choice of the other player was always revealed, it was possible for the participants to learn what the trust outcome would have been even if they decided not to trust the other. Thus, if the participant chose not to trust that could result in two counterfactual outcomes; either the second player would have reciprocated trust or she would have defected trust and taken all the money. As a result, all participants (even those that never trusted) received exactly the same information about the other players' decisions to share or not during the experiment.

The participants were told that the other player made his or her decisions through an Internet connection but in reality the choice was made by the computer program and was displayed after a variable delay of 2–4 seconds. The presentation of this decision of the other player was displayed with an arrow by the outcome of choice. Blue arrows indicated a real outcome following a trust decision; grey arrows indicated a counterfactual outcome following a no trust decision. The presentation of the outcome of the trial was displayed for 3 seconds.

Participants were informed that during the experiment they were playing with three other unknown players (and would be paired with one of those three each round). However, they actually played with computer-simulated agents with different preprogrammed strategies. The players were programmed with different percentages of sharing choices (trustworthy: 80%, neutral: 50%

and untrustworthy: 20%). To represent the other players we used photographs of participants of the same age and gender. Prior to the experiment, the pictures were judged independently by eight students on trustworthiness. Based on those judgments the most neutral faces on the trust dimension were selected for the experiment. To ensure that the individual characteristics of the faces did not bias trusting behavior we randomized the different faces over the different strategies. In total, the task consisted of a single session with 30 interactions with the three computer players. Consequently, for each participant the task consisted of 90 rounds in total. Participants were told they would play for several rounds but, to avoid end-game effects, they were unaware of the exact number of trials. In each round the computer randomly picked one of the three other players, and the total number of rounds was unknown to the participants. The experiment was self-paced and took about 15 minutes to complete.

Finally, the participants were told that the money they earned in the game would be exchanged for real money they would receive at the end of the experiment. We did not mention what the exact exchange rate between game and real money would be, but emphasized that the more money they earned the higher their real payoff would be. Furthermore, the participants were told that their personal income would be revealed only when all other participants finished the experiment.

Postgame questionnaire

Right after the last round of the STG, the participants filled in a computer-based questionnaire. This questionnaire was not mentioned to the participants before they played the STG in order not influence their behavior in the game. We asked three questions regarding the frequency of sharing decisions of the other, level of trustworthiness, and feelings of anger. The first three questions could be answered on a 5-point scale, ranging from *not at all* to *very* (often). We asked the participants to indicate

their estimations of the frequency of sharing and levels of trustworthiness of the other players in order to check whether the different age groups have a comparable perception of how the other players behaved during the game, and how perceived behavior of the others is related to perceived trustworthiness of those players.

Finally, the participants had the opportunity to reduce the earnings of the other players. However, this reduction was costly; for each coin (€0.50) paid by the participants the other player would lose 3 coins (€1.50; cf. Fehr, 2002). For each of the other players the participant could choose to pay an amount between €0.00 and €2.00 in increments of €0.50. The order of the presentation of the three other players was randomized across participants.

Procedure

Child and adolescent participants were individually tested at their school in a quiet room and adult participants were tested in a laboratory, using a standard desktop computer or a laptop. Before the experiment started all participants received verbal instructions and filled out a questionnaire to test whether they understood the structure of the game. Subsequently, they played 10 practice rounds to get familiar with the interface. In case participants made mistakes in the questionnaire, the experimenter personally went over the questions with the participant to verify instructions were understood and if they were not correct they would go through another set of practice rounds until the task was fully understood.

Instructions

All participants got their picture taken a week before they participated in the experiment, and were told their picture would be shown the other players they interacted with in the experiment that would follow. The participants were instructed that they were going to play an interactive game with two other players with whom they were connected via the Internet. Furthermore,

they were told that at the end of the experiment the computer would determine the payoff for all players. It was emphasized that therefore their decisions had consequences for the payoff of themselves *and* others. The total duration of the experiment was approximately 35 minutes. Last, when all participants had completed the experiment, all participants were paid €3.00 and debriefed about the actual setup.

Results

First we tested whether the different age groups differed in their perceptions of frequency of sharing and trustworthiness. Next we investigated how participants of different ages learned whom to trust and distrust, and analyzed the relation between age, anger, and punishment.

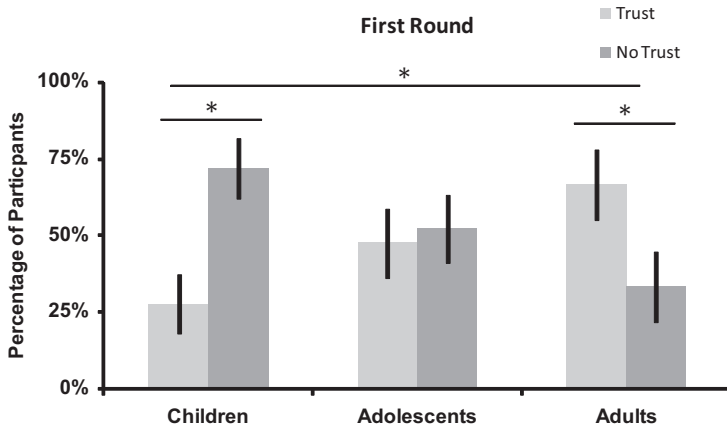
Manipulation check

To check whether there were age differences in the perception of the frequency of sharing decisions of the three types of players, we performed ANOVA with frequency of reciprocal choices as dependent variable, type of player as within-subjects variable and age group as between-subjects factor. These analyses revealed only a main effect of type ($F(2, 58) = 163.20, p < .001$); participants of all age groups recognized that the three players differed significantly in their frequency of sharing decisions (see Table 1), and frequency estimations did not differ between age groups. These results are important because they show that age differences in punishing behavior or emotions are not due to different perceptions of the strategies of the other players.

To investigate whether the different strategies of the other players were correctly recognized as differences in trustworthiness we performed a similar ANOVA with trustworthiness as dependent variable. These analyses revealed only a main effect of type ($F(2, 58) = 138.22, p < .001$). That is, participants of all age groups perceived the three players differing significantly in their trustworthiness (see Table 1), but importantly these estimations did not differ between age groups.

Table 1. Average Levels of Frequency Estimation and Trustworthiness Collapsed Over All Age Groups

	Trustworthy	Neutral	Untrustworthy
Frequency	4.46	2.93	1.63
Trust	4.03	2.72	1.70

**Figure 2.** The percentage of participants in each group that decided to either trust and not to trust in the very first round of the experiment. Error bars represent standard error.

General trust: The first move

As expected, our data show that 11 year olds made fewer trust decisions ($M = 27\%$) in the first round relative to the 16 year olds ($M = 47\%$) and the 22 year olds ($M = 70\%$) who trusted the most (see Figure 2). Indeed, a binary logistic regression with first choice as dependent variable and age group as covariate revealed that with increasing age participants showed significantly more trust in the first round ($\beta = .90, p < .01$; see Figure 2).

Relation-specific changes in trust

Next, we were interested in how trust relations changed over time based on the behavior of the other player, and whether there were age differences in these developing patterns of trust. To investigate the relation-specific changes in trust over time we divided the experiment in three

equal blocks (begin, middle, end). We performed a repeated-measures ANOVA with type (trustworthy, neutral, untrustworthy) and time (begin, middle, end) as within-participants factors and age as between-participants factor for the percentage of trust choices.

As expected, this analysis yielded a main effect of type on trust ($F(2, 58) = 128.03, p < .001$); each player trusted the most trustworthy player the most and the least trustworthy the least. This main effect was qualified by a significant Type \times Time interaction ($F(4, 58) = 28.98, p < .001$); over time participants showed *increasing* trust for the trustworthy player and *decreasing* trust for the untrustworthy player (see Figure 3). Moreover, our analyses also revealed a Type \times Time \times Age interaction ($F(4, 58) = 13.14, p < .005$). This indicates that there are age differences in relation-specific changes in trust, as can be seen in Figure 3.

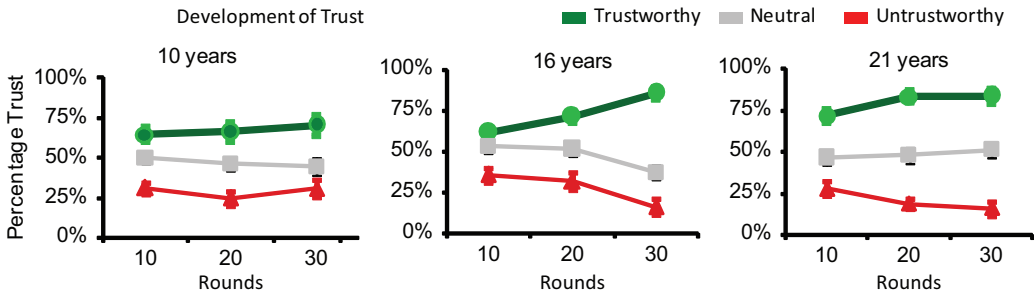


Figure 3. The mean percentage of trust decisions per block of 10 trials for both computer players, error bars represent standard error.

To further interpret these age differences we performed separate Type \times Time ANOVAs for each age group. These analyses revealed that in all age groups there was a significant difference in the amount of trust in each of three players (i.e., a main effect of type, all p s $< .001$). Furthermore, for adults and adolescents ($F(4, 18) = 16.11, p < .001$ and $F(4, 19) = 8.74, p < .005$ respectively) but not for the children ($F(4, 18) = 2.67, p = .08$) we observed a significant Type \times Time interaction. The pattern of the children differs from the other age groups by showing no significant change in strategy over time, whereas adults and adolescents started to trust the trustworthy player more, and the untrustworthy player less, over time (see Figure 3).

Positive versus negative reciprocity

To investigate differences in strategy use during the game we analyzed sequential effects. We analyzed whether the choice of the other player to reciprocate or defect in the previous round, regardless of whether the outcome was real or counterfactual (i.e., following trust or no trust), influenced the participants' decision to trust in the next round with the same player. For this purpose we distinguished between positive reciprocity (i.e., making a decision to trust after the other player decided to share on the previous round) and negative reciprocity (i.e., making a decision not to trust after the other player decided to not share on the previous round). To investigate

possible differences in reciprocal choices after either a share or keep choice of the other we performed an ANOVA with reciprocal choice type (positive vs. negative) as within-participants factors and age as between-participants factor. This analysis revealed group differences in reciprocal choices (Choice Type \times Age interaction, $F(2, 58) = 4.67, p < .01$). Post hoc ANOVA analyses showed developmental change in negative ($F(2, 58) = 7.05, p < .002$), but not positive reciprocity, ($F(2, 58) < 1, p = .59$). Additionally, post hoc paired t tests per age group shows that children showed higher levels of negative than positive reciprocity ($t(1,17) = -2.46, p < .025$), whereas the other groups did not show such a difference (both p s $> .3$, see Figure 4). These results indicate that the youngest group is more sensitive to trust violations and that this sensitivity declines across adolescence.

Anger and punishment

Next, we investigated how the different strategies used by the computer players elicited feelings of anger and subsequent punishment in the three different age groups. We performed an ANOVA with anger as dependent variable, type of player as within-subjects variable and age group as between-subjects factor. This analysis revealed only main effects of type ($F(2, 58) = 12.70, p < .001$), and age ($F(2, 58) = 12.69, p < .001$). That is, participants of all age groups showed more anger to the least trustworthy person; and as expected

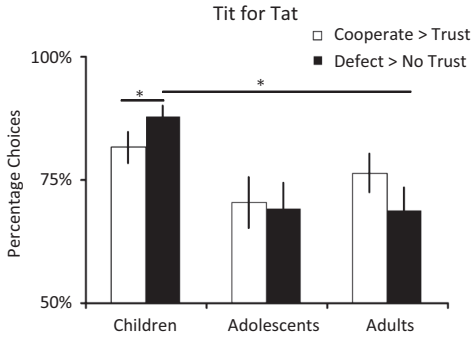


Figure 4. The mean percentage of negative and positive reciprocity decisions for all age groups.

the younger participants also showed more anger than the older participants (see Figure 5).

The analysis of punishment behavior revealed a similar pattern to that of anger; participants of all age groups punished the least trustworthy person the most (main effect type, $F(2, 58) = 16.12, p < .001$), and with age there was a general decrease in the amount of punishment given (main effect age, $F(2, 58) = 5.08, p < .03$). Finally, we were interested in the relation between reported levels of anger and subsequent size of punishment. As expected, there was a significant correlation between anger and punishment for all age groups ($r = .54, p < .01, r = .53, p < .01$ and $r = .42, p < .03$ for children, adolescents, and adults). The correlation between anger and punishment suggests that the decrease in punishment with age is a result of decreased anger with age. To further investigate our hypothesis that developmental changes in trust behavior are related to changes in emotion regulation we performed an additional mediation analysis.

Mediation can be demonstrated by showing that the indirect effect (i.e., the path from age to punishment through the mediator anger) is significantly different from zero. The indirect effect is the product of two regression coefficients; specifically, the product of the regression weight linking the independent variable to the mediator (denoted a) and weight linking the mediator to the dependent variable (denoted b , see Figure 6). Shrout and Bolger (2002) suggest that a formal

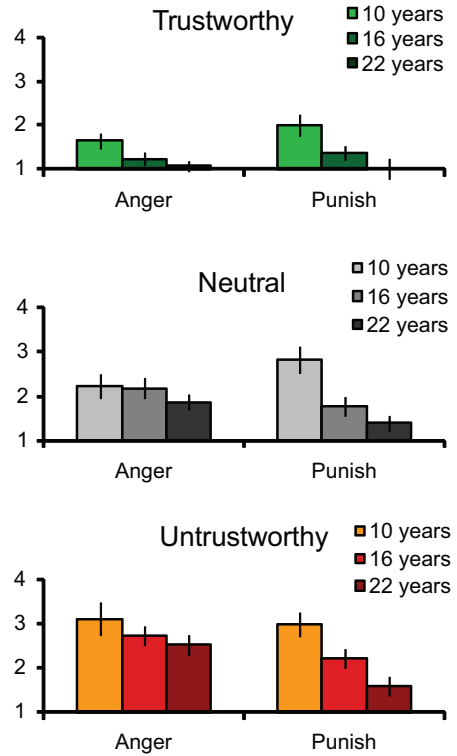


Figure 5. Post-Game Questionnaire results for anger and punishment.

test of mediation be conducted using a bootstrapping technique that involves computing confidence intervals around the product term ($a*b$). If zero falls outside of this 95% confidence interval, the indirect effect is significant and mediation can be said to have occurred. To implement this approach, we used SPSS syntax provided by Preacher and Hayes (2004) using 10,000 iterations. The results indicate that zero fell outside our 95% confidence interval around the indirect effect, which ranged from .17 to .54. Thus, these results provide evidence that anger mediates the effects of age on punishment.

Discussion

Despite strong evidence for the benefits of trust for social development and society (e.g., Bernath & Feshbach, 1995; Fukuyama, 1995), it is less

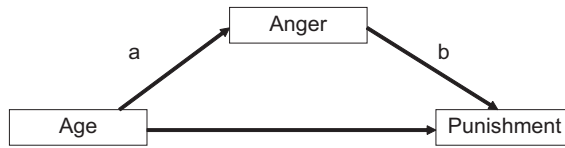


Figure 6. Mediation diagram of age, anger and punishment.

well known how people learn to trust or distrust persons and anonymous institutions. In this article, we combined insights from social psychology and developmental psychology, and used the STG to study the relation-specific changes in trust in three age groups. To our knowledge, no study to date has investigated age differences in learning whom to trust and costly punishment of trust violations. This study had two main goals: (a) to examine the development of trust relationships between late childhood and young adulthood, and (b) to examine the developmental trajectory of emotions evoked by noncooperative behavior of others, and to what extent these emotions may lead to altruistic punishment. To this end, the discussion is organized according to these main goals.

Learning whom to trust

As noted, previous research with the trust game has paid some attention to trust in children, but almost without exception these studies involved adults only. The decisions of adults in the current study resemble the pattern typically seen in these behavioral experiments. That is, adult participants often chose to trust in the first round, indicating that they expected others to reciprocate (e.g., Berg et al., 1995; Dufwenberg & Gneezy, 2000; McCabe et al., 2001). However, there were important age-related changes in first move. As expected, only a small number of children trusted on the first move, and the number of trust choices in the first round increased gradually with age. These results are consistent with previous studies with single-round trust games that showed that general trust increases with age (Harbaugh et al., 2002; Sutter & Kocher, 2008). Note that the first trusting move may to some

extent reflect strategic self-presentation, done to elicit cooperative responding from a partner (Danheiser & Graziano, 1982). Interestingly, this strategic self-presentation depends on the ability of strategically taking the perspective of the other person; a skill that still develops during this period (Dumontheil, Apperly, & Blakemore, 2010; van den Bos et al., 2011). It might therefore be possible that perspective taking also plays a role in multiple-round trust games. Future studies are needed to further investigate the role of perspective taking in multiple-round trust games. Additionally, in order to strengthen construct validity, future studies could benefit from additional measures of dispositional trust and perspective-taking skills besides in-game decisions.

Although the low level of general trust displayed by children in the first trial is consistent with previous studies, the following question remained: how would children and adolescents learn to trust or distrust another player? Our analyses of the relation-specific changes in trust revealed that participants of all ages were able to learn to trust a certain player, and importantly also learn not to trust another player. Indeed, both children and adolescents ended with high levels of trust for the trustworthy player and low levels of trust for the untrustworthy player. Interestingly, there were also age differences in strategies.

As expected we observed significant increases and decreases in levels of trust for both adolescents and adults, indicating that they were updating their initial level of trust based on the positive and negative outcomes they experience during the task (Camerer & Ho, 1999; Erev & Roth, 1998). In contrast, although children differentiated between the three players,

their levels of trust did not change as much during the experiment. Additionally, further analyses revealed that children differed from adults and adolescents in showing higher levels of negative reciprocity. Taken together, the data suggest that during development the level of trust becomes; (a) more dependent on the total history of interactions instead of just the most recent ones, and (b) becomes more resistant to violations of trust. These results support our hypothesis that children are less able to regulate the anger evoked by violations of trust. As such, the results are consistent with studies that have shown that children are less capable to regulate their emotions in social situations than adults (Eisenberg, 2000), and studies that have shown that age-related increases in emotion regulation are strongly related to lower levels of reactive aggression (Connor et al., 2004; Winstok, 2009). In this case children are reacting more strongly to violations of trust in the previous round, regardless of the other players' indications of trust behavior in the past. This interpretation is further supported by our analyses of the post-game questionnaire.

Anger and punishment

We investigated participants' emotional reactions to trust violations and levels of costly punishment. As expected, the three players evoked different levels of both anger and punishment. Participants of all age groups were most angry at the player that violated trust the most, and punished accordingly. This pattern of behavior is consistent with several previous studies that investigated the relation between anger and costly punishment (see Seip et al., 2009). However, there were also large differences in levels of anger between age groups. Although all participants displayed more anger towards those players that violated trust the most, children showed more anger than adolescents, and adolescents more than adults. Furthermore, in line with our hypothesis we found that the younger participants punished more than the older participants.

In contrast to children, adults showed virtually no anger towards, and did not punish, the least untrustworthy player, even though that player kept the money 20% of the time. So, although that player displayed some trust violations, these occasional violations did not seem to anger the adult participants, and it did not induce them to punish. Interestingly, the mediation analysis did suggest that the age-related changes in punishment are mediated by individual levels of anger towards the participants, supporting the hypothesis that emotion regulation is an important factor in driving (developmental) differences in trust behavior.

Finally, an additional reason for the higher levels of anger in children is that their affective reaction to social interaction is based on a different perception of the intentions of the other players. Previous developmental studies have suggested that the increased skill of perspective taking, the ability to reason about the others' intentions, significantly changes social behavior in one-shot trust (van den Bos et al., 2009) and ultimatum games (Guroğlu et al., 2009; Sutter, 2007). These studies suggest that an age-related increase in perspective taking may lead to increased trust and a decrease in rejection rates. Furthermore, Mohr, Howells, Gerace, Day, and Wharton (2007) showed that increased anger after provocation (i.e., violation of trust) is significantly related to a decreased capability of perspective taking (see also Eisenberg & Fabes, 1998). Taken together, this suggests that a possibly more negative perception of the others' intentions by the younger participants could have led to more anger and subsequently more punishment after the violation of trust. Given that all age groups had similar perceptions of the trustworthiness of the three players, the current results favor the explanation of differences in emotion regulation over perspective taking. Future studies may focus on disentangling the effects of perspective taking and emotion regulation on the increased negative affect in developmental populations, which will further our understanding of these processes in social decision making.

Conclusion

The current findings revealed the importance of several psychological processes involved in learning whom to trust. A comparison of age differences of behavior in the STG indicates that, besides a general increase of generalized trust, relation-specific trust changes with age. In particular, children appeared to be especially more sensitive to violations of trust. Additionally, the results show that with increasing age the amount of both anger and punishment decreased, and that age differences in trust were fully mediated by feelings of anger. Together these support the hypothesis that the stability of adult trust relationships might be the result of an age-related increase in the ability to regulate negative affect associated with violations of trust, resulting in less negative reciprocity and lower punishment.

Finally, the current findings demonstrate how the combination and integration of social psychological and developmental insights may contribute to understanding of how we learn to trust (and distrust) others. The results extend the literature on learning in context of the trust game by showing the importance of affect and affect regulation in the development of trust over multiple interactions. Recently it has been argued that affect plays an important role in coordinating group processes (Kelly, 2001; Spoor & Kelly, 2004). One interesting avenue to explore the role of affect in more detail would therefore be to study and compare the development of trust within and between groups (Cox, 2002; Kugler, Bornstein, Kocher, & Sutter, 2007).

Note

1. To test for possible effects of gender, we also performed our behavior analyses controlling for gender of the participant. The results of those analyses did not differ from the currently reported results.

Acknowledgments

The authors would like to thank Sema Adiguzel, Yoricka Paasen, Lizette Romijn, and Zeynep Samur for their help with the data collection. This research was supported by a Vidi grant from The Netherlands Organisation for Scientific Research to Eveline A. Crone.

References

- Barney, J. B., & Hansen, M. H. (1994). Trustworthiness as a source of competitive advantage. *Strategic Management Journal*, *15*, 175–190.
- 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.
- Boon, S. D., & Holmes, J. G. (1991). The dynamics of interpersonal trust: Resolving uncertainty in the face of risk. In R. A. Hinde & J. Groebel (Eds.), *Cooperation and prosocial behavior* (pp. 167–182). New York, NY: Cambridge University Press.
- Boyle, R., & Bonacich, P. (1970). The development of trust and mistrust in mixed-motives games. *Sociometry*, *33*, 123–139.
- Camerer, C. F., & Ho, T. (1999). Experience-weighted attraction in games. *Econometrica*, *64*, 827–874.
- Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, *1124*(1), 111–126.
- Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences*, *26*, 139–198.
- Connor, D. F., Steingard, R. J., Cunningham, J. A., Anderson, J. J., & Melloni, R. H., Jr. (2004). Proactive and reactive aggression in referred children and adolescents. *American Journal of Orthopsychiatry*, *74*, 129–136.
- Cox, J. C. (2002). Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. In R. Zwick & A. Rapoport (Eds.), *Advances in experimental business research*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Danheiser, P. R., & Graziano, W. G. (1982). Self-monitoring and cooperation as a self-presentational strategy. *Journal of Personality and Social Psychology*, *42*(3), 497–505.
- 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., & Coie, J. D. (1987). Social-information processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology*, *53*, 1146–1158.
- Dufwenberg, M., & Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior*, *30*, 163–182.

- Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science, 13*(2), 331–338.
- Dunn, J. R., & Schweitzer, M. E. (2005). Feeling and believing: The influence of emotion on trust. *Journal of Personality and Social Psychology, 88*(5), 736–748.
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology, 51*, 665–697.
- Eisenberg, N., & Fabes, R. A. (1998). Prosocial development. In W. Damon (Ed.), *Handbook of child psychology. Vol. 3: Social, emotional, and personality development* (pp. 701–778). New York, NY: Wiley.
- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in games with unique strategy equilibrium. *American Economic Review, 88*, 848–881.
- Fehr, E. (2002). The nature of human altruism. *Nature, 415*, 269–272.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*, 785–791.
- Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. New York, NY: Free Press.
- Gouldner, A. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review, 25*, 161–178.
- Gulati, R. (1995). Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances. *The Academy of Management Journal, 38*(1), 85–112.
- Güroğlu, B., van den Bos, W., & Crone, E. A. (2009). Fairness considerations: Increasing understanding of intentionality in adolescence. *Journal of Experimental Child Psychology, 104*, 398–409.
- Harbaugh, W. T., Krause, K., Liday, S. G., & Vesterlund, L. (2002). Trust in children. In E. Ostrom & J. Walker (Eds.), *Trust, reciprocity and gains from association: Interdisciplinary lessons from experimental research* (pp. 302–332). New York, NY: Sage.
- Huang, L., & Murnighan, J. K. (2010). What's in a name? Subliminally activating trusting behavior. *Organizational Behavior and Human Decision Processes, 111*(1), 62–70.
- Imber, S. (1973). Relationship of trust to academic performance. *Journal of Personality and Social Psychology, 28*, 145–150.
- Kahn, E. H., & Turiel, E. (1988). Children's conceptions of trust in the context of social expectations. *Merrill-Palmer Quarterly, 34*, 403–419.
- Kelly, J. R. (2001). Mood and emotion in groups. In M. A. Hogg & R. S. Tindale (Eds.), *The Blackwell handbook of social psychology, Vol. 3: Group processes* (pp. 164–181). Oxford, UK: Blackwell.
- King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., & Montague, P. R. (2008). The rupture and repair of cooperation in borderline personality disorder. *Science, 321*, 806–810.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science, 308*, 78–83.
- Kugler, T., Bornstein, G., Kocher, M. G., & Sutter, M. (2007). Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy. *Journal of Economic Psychology, 28*(6), 646–657.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review, 20*(3), 709–734.
- McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal, 38*, 24–59.
- McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings National Academy Science United States of America, 98*(20), 11832–11835.
- Mohr, P., Howells, K., Gerace, A., Day, A., & Wharton, M. (2007). The role of perspective taking in anger arousal. *Personality and Individual Differences, 43*, 507–517.
- Murnighan, J. K., & Saxon, M. S. (1998). Ultimatum bargaining by children and adults. *Journal of Economic Psychology, 19*, 415–445.
- Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavioral Human Decision Processes, 68*, 208–224.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, and Computers, 36*, 717–731.
- Rotenberg, K. J. (1980). "A promise kept, a promise broken": Developmental bases of trust. *Child Development, 51*, 614–617.
- Rotenberg, K. J., & Pilipenko, T. A. (1983–1984). Mutuality, temporal consistency, and helpfulness in children's trust in peers. *Social Cognition, 2*, 235–255.

- Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, *53*, 651–655.
- Rotter, J. B. (1971). Generalized expectancies for interpersonal trust. *American Psychologist*, *26*, 443–452.
- Seip, E., van Dijk, W. W., & Rotteveel, M. (2009). On hotheads and Dirty Harries: The primacy of anger in altruistic punishment. *Annals of the New York Academy of Sciences*, *1167*, 190–196.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422–445.
- Spoor, J. R., & Kelly, J. R. (2004). The evolutionary significance of affect in groups: Communication and group bonding. *Group Processes & Intergroup Relations*, *7*(4), 398–412.
- Stouten, J., De Cremer, D., & van Dijk, E. (2009). Behavioral (in)tolerance of equality violation in social dilemmas: When trust affects contribution decisions after violations of equality. *Group Processes and Intergroup Relations*, *12*, 517–531.
- Sullivan, J. L., & Transue, J. E. (1999). The psychological underpinnings of democracy: A selective review of research on political tolerance, interpersonal trust, and social capital. *Annual Review of Psychology*, *50*, 625–650.
- Sutter, M. (2007). Outcomes versus intentions: On the nature of fair behavior and its development with age. *Journal of Economic Psychology*, *28*, 69–78.
- Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior*, *59*(2), 364–382.
- Van den Bos, W., van Dijk, E., Westenberg, P. M., Rombouts, S. A.R. B., & Crone, E. A. (2011). Changing brains, changing perspectives. *Psychological Science*, *22*(1), 60–70.
- Van den Bos, W., Westenberg, P. M., van Dijk, E., & Crone, E. A. (2010). Development of trust and reciprocity in adolescence. *Cognitive Development*, *25*(1), 90–102.
- Wentzel, K. R. (1991). Relations between social competence and academic achievement in early adolescence. *Child Development*, *62*(5), 1066–1078.
- Winstok, Z. (2009). From self-control capabilities and the need to control others to proactive and reactive aggression among adolescents. *Journal of Adolescence*, *32*(3), 455–466.
- Wright, T. L., & Kirmani, A. (1977). Interpersonal trust, trustworthiness and shoplifting in high school. *Psychological Reports*, *41*, 1165–1166.
- Yamagishi, T., Cook, K. S., & Watabe, M. (1998). Uncertainty, trust, and commitment formation in the United States and Japan. *American Journal of Sociology*, *104*, 165–194.
- Zak, S., & Knack, P. (2001). Trust and growth. *Economic Journal*, *111*, 295–321.