Journal of Trust Research

ISSN: 2151-5581 (Print) 2151-559X (Online) Journal homepage: http://www.tandfonline.com/loi/rjtr20

Deterrence-based trust in bargaining: Introducing a new experimental paradigm

Eric van Dijk, Varia Makagonova, Erik W. de Kwaadsteniet & Manon Schutter

To cite this article: Eric van Dijk, Varia Makagonova, Erik W. de Kwaadsteniet & Manon Schutter (2017) Deterrence-based trust in bargaining: Introducing a new experimental paradigm, Journal of Trust Research, 7:1, 71-89, DOI: 10.1080/21515581.2016.1254093

To link to this article: https://doi.org/10.1080/21515581.2016.1254093

JOURNAL OF TRUST RESEARCH, 2017
VOL. 7, NO. 1, 71–89
http://dx.doi.org/10.1080/21515581.2016.1254093

Routledge
Taylor & Francis Group

Deterrence-based trust in bargaining: Introducing a new experimental paradigm

Eric van Dijk, Varia Makagonova, Erik W. de Kwaadsteniet and Manon Schutter

Department of Social and Organizational Psychology, Institute for Psychological Research, Leiden University, Leiden, The Netherlands

ABSTRACT

Trust, especially in the initial stages of trust building, is often assumed to be the result of deterrence-based trust. While theorising acknowledges its importance, research on deterrence-based trust has been scarce. To facilitate the investigation of the concept, we designed new versions of the trust game in which we studied both trust (Experiment 1) and trustworthiness (Experiment 2). To better model deterrence-based trust we extended a trust game with an additional phase where trustors could accept or reject the trustee's distribution. We varied consequences of the rejection option as a delta bargaining game, thereby manipulating the potential for deterrence. The results showed that trustors were highly responsive to the possibility to reject the trustee's distribution. Trustees, however, seemed largely unaffected and were generally highly trustworthy. Together these findings show how trust games can meaningfully be extended to assess the effect of deterrence-based trust in bargaining.

ARTICLE HISTORY
Received 1 February 2016
Accepted 22 September 2016

KEYWORDS
Deterrence-based trust; distrust; trustworthiness; bargaining; games

ACTION EDITOR
Kong, Dejun Tony

Trust and vulnerability tend to go hand in hand. The essence of trust is that you willingly put yourself in a position in which others may take advantage of you. Indeed, it is this aspect that was central in Rousseau, Sitkin, Burt, and Camerer's (1998, p. 395) definition of trust as 'a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behaviors of another'.

Trust can be based on positive expectations about the good-will or benevolence of others (Mayer, Davis, & Schoorman, 1995). Positive expectations can, however, also have a more situational component (see for an elaborate discussion of personal vs. situational bases of trust e.g. Thielmann & Hilbig, 2015a). Deterrence-based trust can be defined as the trust people have in others when they believe that for these others the costs for breaches of trust will outweigh the benefits of untrustworthiness. Related to this, Lewicki and Bunker (1996) described the notion of calculus-based trust as a broader concept by acknowledging that trust may not only be based on the presence of negative consequences of breaches of trust for others (i.e. if they can be punished),

CONTACT Eric van Dijk dijk@fsw.leidenuniv.nl Department of Social and Organizational Psychology, Institute for Psychological Research, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands

but also on the positive consequences of trustworthy behaviour (i.e. if they can be rewarded).

Shapiro, Sheppard, and Cheraskin (1992) referred to deterrence-based trust as the most basic form of trust that is especially relevant in the initial stages of trust building (see also Lewicki, Tomlinson, & Gillespie, 2006; Rousseau et al., 1998). The concept of deterrence-based trust – and its more general form of calculus-based trust – adds a utilitarian underpinning to one's expectations (see Rousseau et al., 1998; Shapiro et al., 1992) that is not based on some inherent positive evaluation about others, but rather on the consideration that others will refrain from untrustworthy behaviours to the extent that it is too costly for them to do so.[1] If the trustees' costs are dependent on the trustor's behaviour (e.g. if the trustor is the one who can punish the trustee), a level of interdependence is created that adds a strategic element to the setting. Lewicki and Polin (2013) deemed this aspect of interdependence of special relevance to bargaining and other settings with repeated interaction.

While these insights all point to the relevance of deterrence-based trust, it should be acknowledged that this concept has rarely been the subject of systematic research. That is, insights on the role of deterrence-based or calculus-based trust have often stayed at the theoretical level. Ten years ago, when reviewing theory and research on interpersonal trust, Lewicki et al. (2006, p. 992) stated that 'a number of definitions and conceptualisations have been proposed, yet efforts to measure trust and its component elements have not kept pace'. Today, this conclusion still holds; definitely when it comes to studies on bargaining. The scarce research that has assessed the importance of these types of trust used questionnaires to tap the concept. For example, in their study on three-party negotiations, Olekalns, Lau, and Smith (2007) assessed calculus-based trust with items like 'This person will know that the benefits of maintaining trust are higher than the costs of destroying it.' In a similar vein, Olekalns, Kulik, and Chew (2014) assessed deterrence-based trust with items like 'If this person doesn't do what he/she is going to say, I can get even.' While these studies can be considered as an advancement and very welcome response to Lewicki et al.'s (2006) statement, we now aim to further advance the research on deterrence-based trust by designing a new paradigm that allows for behavioural measures of deterrence-based trust.

A behavioural measure of deterrence-based trust

Questionnaires to assess people's trust, and in this case their deterrence-based trust, basically see it as an internal state on which respondents are asked to report. While such assessments are certainly helpful, it should be noted that such reports do have their limitations, and are often only weakly predictive of actual behaviour (see e.g. Glaeser, Laibson, Scheinkman, & Soutter, 2000). To address such limitations, research has also developed behavioural measures of trust. The most prominent example is the behaviour that is measured in the Trust Game, developed by Berg, Dickhaut, and McCabe (1995), which since its publication has further stimulated the research on trust. The Trust Game models the situation in which individuals may obtain higher outcomes if (a) they trust others, and (b) these others prove themselves trustworthy. It depicts an experimental setting with two players, who decide one after the other. Player A (the trustor) is usually endowed with a certain amount of money (e.g. $10) which he/she can keep to him/herself. However, player A can also decide to let player B (the trustee) distribute the

money. In that case, this transferred amount is tripled before player B makes the distribution. So, if A would transfer 1 of the $10 to B, B could distribute $3; when allocating $4, B could distribute $12, etcetera. In its original form, player A could make a continuous decision and decide to allocate any amount between the 0 and $10 (see for a review on behaviour in trust games with continuous decisions Johnson & Mislin, 2011). Other versions have been used as well in which A's decision, for example, was binary, such that A could only to decide to allocate nothing (no trust) or all outcomes to B (trust; see for binary versions of trust games e.g. Bohnet & Zeckhauser, 2004; Fetchenhauer & Dunning, 2009; Kreps, 1990; Malhotra, 2004; Schilke, Reimann, & Cook, 2015).

It is clear that – both in the continuous and the binary version – there is a potential benefit for both players to let B divide the (then tripled amount of) money, that is, to trust B. However, the risk is on A because player B may prove him/herself untrustworthy and take the bulk of money and run, leaving A with little or no money. Trust games nicely capture the element of (A's) vulnerability, where it only makes sense for A to hand the money to B when having positive expectations about B's willingness to return the favour (e.g. by returning half of the tripled money). Trust in this game thus captures 'behavioural trust', or 'trust as a choice' which is often perceived as more telling than 'trust as an attitude' (see also Li, 2007, 2012). Note, however, that the typical trust game as described above also has its limitations in the sense that A has no other option than to accept B's distribution (even if B allocates nothing to A). Put differently, B is in a position where he/she can react to A's behaviour, but A can never respond to B's behaviour, let alone retaliate.

In this sense, the typical trust game is not yet suited to capture deterrence-based trust. One could thus also claim that while the typical trust game does capture trust and trustworthiness, it only captures part of it. In bargaining settings, interdependence generally has more of a back and forth nature, where you could, for example, reject an offer. Having an option to turn down B's allocation would not only add the element of further bargaining, but provide A with the additional means of being able to make untrustworthy behaviour costly. In other words, it would add an element of deterrence-based trust to the trust game because betrayal of trust may now come at a cost. It is this aspect of interdependence that we intended to incorporate, and which allows us to test its separate effect on trust and trustworthiness.

The most direct and straightforward way to incorporate this in a trust game setting is to simply provide A with the option to reject B's distribution. In terms of behavioural games, this means replacing B's final distribution – that cannot be rejected in the typical trust game – with a distribution that *can* be rejected, with costly consequences for the trustee. In the current study, we did this by replacing it with a bargaining game that is modelled after the ultimatum bargaining game (Güth, Schmittberger, & Schwarze, 1982). In the ultimatum bargaining game, one player makes an offer to the other player who can either accept or reject it; if the offer is rejected, both players obtain nothing. Added to the trust game setting, this thus means that if A is not satisfied with B's distribution of the tripled money, he/she can still reject the offer in which case neither player receives anything. While this may not be an attractive prospect for A either, it does provide A with the behavioural option of deterrence; that is, with a means to incur costs on B.

That such options may work to prevent exploitation has been demonstrated in research using variants of the ultimatum game, formally described as the 'delta game' (Suleiman,

1996). The delta game is a more general form of the ultimatum game in that upon rejection, the rejected offer is multiplied by a factor delta ($0 \leq \text{delta} \leq 1$). When delta equals zero, you indeed have the ultimatum game where rejection of the offer results in zero outcomes for both players. If delta equals 1, the game turns into a dictator game (e.g. Bolton, Katok, & Zwick, 1998) where the distribution cannot be changed even if you reject it. An in-between value, delta = 0.5, means that upon rejection, the offer is reduced by 50%. Generally, research on the delta game has revealed that offers go up as delta decreases (Handgraaf, Van Dijk, Vermunt, Wilke, & De Dreu, 2008; Suleiman, 1996). For example, offers are generally higher if delta = 0 (the traditional ultimatum game) than if delta = 0.5, the main explanation being that with lower levels of delta, people fear the consequences of making too low offers. Low deltas thus seem to function as a safeguard against exploitation.

If we combine these insights with the notion of deterrence-based trust, and incorporate the delta game into the traditional trust game setting, we come to the 'Extended Trust Game' format depicted in Figure 1. The essential change is captured in the grey box. Whereas in the traditional trust game, after handing the decision to player B, players A have no other option than to accept B's distribution, the extended version we designed here allows for rejection of this distribution. By doing so, the trust game transfers to a setting that enables us to behaviourally measure deterrence-based trust.

So how would this affect the trust process? How would trust and trustworthiness fare if we extend player A's possibilities for reciprocation by providing him/her with the possibility to accept/reject B's distribution? To answer these questions, a distinction has to be made between the two parties involved. For player A, it seems plausible to assume that by adding deterrence-based trust (i.e. with an opportunity to reject B's distribution), A will be more likely to behaviourally trust B. The option to reject may increase deterrence-based trust, and thus increase A's willingness to let B divide the (tripled) amount, especially if the consequences of rejection are high for B (cf. Suleiman, 1996).

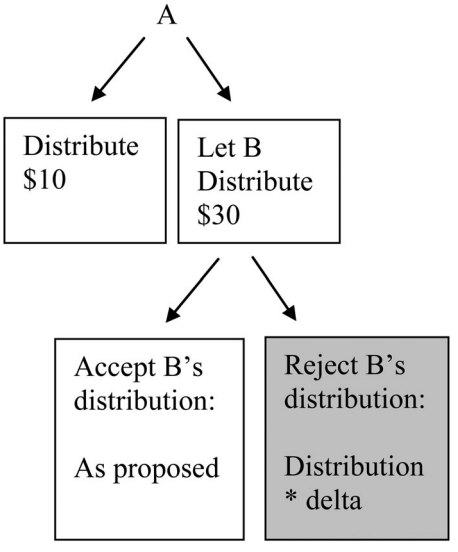

Figure 1. The extended trust game.

It is less clear, however, how it would affect players B. Based on Suleiman (1996), one might expect that the B's allocation to A will increase when the negative consequences for B of rejection are high. Note, however, that previous research has indicated that players B are generally already quite trustworthy and benevolent; even in the traditional trust game where B could do whatever he/she pleases. This has, for example, been shown in research by Fetchenhauer and Dunning (2009, 2010) who noted that players B are often more trustworthy than assumed by players A. Several reasons may account for this positive (and unanticipated) behaviour, in addition to the fact that people are often simply nicer or benevolent than one would expect. First of all, research on the equal division rule has shown how strong and pervasive the norm of equality is (e.g. Messick, 1995; Van Dijk, Wit, Wilke, & De Kwaadsteniet, 2010). In a trust game the social norm to treat A fairly and distribute outcomes equally is even further strengthened by the reciprocity norm, which prescribes that positive acts should by reciprocated with positive behaviour (see also Ashraf, Bohnet, & Piankov, 2006). Taken together, these combined forces may be so strong that they override more utility-based considerations like a self-interested motivation to benefit oneself at the expense of others. If so, the introduction of deterrence-based trust may have more impact on the generally sceptic player A than on the generally trustworthy player B.

Experiment 1: The willingness to trust

In our first study we focused on player A, and thus on the willingness to trust. As in traditional trust games, the participants made a binary decision by choosing between two options: They could either decide to distribute monetary outcomes themselves or let player B decide on the distribution of outcomes, in which case the amount would first be tripled. Different from the traditional trust game, our participants learned that if they would decide to let B divide the (tripled) outcomes, they could subsequently either accept or reject B's distribution. The consequences were modelled after the delta game (Suleiman, 1996). All participants were informed that this information was common knowledge, so that B was aware of these features as well. In the delta = 1 condition (resembling the traditional trust game), participants learned that even if they would reject the distribution, the outcomes would be divided in accordance with B's distribution. In the delta = .5 condition, participants learned that should they reject B's distribution, the outcomes of both players would be reduced by 50%. In the delta = 0 condition, participants learned that should they reject B's distribution, both players would receive zero outcomes. We reasoned that having an option to reject B's distribution – and thus to impose costs on B should the distribution not be to one's liking – would increase deterrence-based trust. As a result, our operationalised prediction was that the willingness to let B divide the outcomes would increase as delta would go down. Thus, compared to the traditional trust game (delta = 1), the willingness to let B divide the outcomes should be higher in case of delta = 0.5 and delta = 0.

Method

Design and participants

The participants, 84 students (50 males, 34 females; mean age = 21.25 years; *SD* = 2.34) at Leiden University, participated voluntarily. Participants were randomly assigned to one of

three conditions (delta = 1, .5, 0). No participants were excluded from the analyses.[2] The number of participants was determined based on a strategy to obtain as many participants in the weeks we could use the research lab. The main dependent measure was the participants' willingness to trust.[3]

Procedure

Participants were invited to the laboratory to participate in a study on decision-making. Upon arrival, they were placed in separate cubicles, each containing a computer connected to a server. This computer was used to present the information and to register the dependent measures. At the start of the instructions, the participants were informed that they were paired with one of the other participants. Members of each dyad would be referred to as Person A and B. All participants learned that they were assigned the letter A.

The options were described without any reference to the words 'trust' or 'game'. All participants learned that they could distribute chips, each worth 10 cents, and that they could choose between two options: (1) They could either distribute 20 chips themselves, or (2) let B distribute 60 chips.[4] In the delta = 1 condition, participants learned that if they opted to let B distribute 60 chips, they could not influence B's distribution. Participants in the delta = .5 learned that if they opted to let B distribute 60 chips, they would subsequently receive B's distribution, which they could then either accept or reject. In the case of rejection, the distribution would be reduced by 50%. In the delta = 0 condition, participants learned that if they would reject B's distribution, both would receive nothing. Participants learned that this information was also known to B.

After this explanation, participants made their decision by indicating whether they wanted to distribute 20 chips themselves or whether they wanted B to distribute 60 chips. Participants who chose option (1) were asked to distribute the 20 chips. Those who chose option (2) were asked how they would have distributed the 20 chips had they opted for option (1).

We also asked the participants about their expectations about B's distribution. Those who had chosen option (2) were asked how many of the 60 chips they expected B to allocate to them. Those who had chosen option (1) were asked how many of the 60 chips they thought B would have allocated to them, had they chosen option (2).

To provide a direct check of our manipulation of delta – meant to influence the potential consequences should A reject the participants' distribution – we asked three questions: 'How much influence did you have if you would disagree with B's distribution?' (1 = not at all, 7 = very much), 'What would be the consequence for you if you would disagree with B's distribution?' (1 = no consequences; 7 = very serious consequences), and 'What would be the consequence for B if you would disagree with B's distribution?' (1 = no consequences; 7 = very serious consequences). These questions were combined to provide a reliable 'perceived consequences of rejection' scale (Cronbach's alpha = .80).

Finally, we also checked for understanding of the main features of the setting (i.e. whether they had been Person A or B, the value of the chips, whether decisions were anonymous, what the exact consequences of rejection would be).[5] After the study ended, participants were thoroughly debriefed and were paid 3.50 Euros. All participants agreed to this procedure.

Results

Manipulation checks

All participants correctly identified their own role (Person A), and that decisions were anonymous. When asked to indicate the value of the chips (3, 5, or 10 cents), all participants correctly reported that chips were worth 10 cents. When asked what, should they decide to let B divide 60 chips, would be the consequences if they were to reject B's distribution (1 = no consequences; 2 = the distribution would be reduced by 50%; 3 = we would both receive zero chips), 81 of the 84 (96%) participants answered this question correctly.[6]

A one-way ANOVA on the constructed perceived consequences scale confirmed these findings by showing a significant effect for condition ($F(2,81) = 56.62$, $p < .001$; $\eta^2 = .58$): The consequences of rejecting B's distribution were deemed less serious in the delta = 1 condition (resembling the traditional trust game; $M = 2.39$) than in the delta = .5 condition ($M = 4.98$; $SD = 1.14$; $t(55) = -8.72$, $p < .001$; Cohen's $d = 2.31$), and the delta = 0 condition ($M = 5.38$; $SD = 1.18$; $t(53) = -9.76$, $p < .001$; Cohen's $d = 2.63$). The difference between the latter two conditions was not significant ($t(54) = -1.31$, $p = .20$; Cohen's $d = 0.35$).

These results indicate that the participants correctly understood the main characteristics of the experiment, and that our manipulations were successful, while it is noteworthy that the participants did not significantly differentiate between perceived consequences of the delta = 0 and the delta = .5 condition. We return to this observation in the Discussion of Experiment 1.

Trust decision

The decision to trust or not was significantly affected by our manipulations ($\chi^2(2) = 7.34$, $p = .026$). Closer inspection revealed that participants in the delta = 1 condition (resembling the traditional trust game) were significantly less willing to let B divide the chips than those in the delta = 0.5 ($\chi^2(1) = 4.10$, $p = .043$) and delta = 0 conditions ($\chi^2(1) = 6.03$, $p = .014$). The number of participants choosing to trust did not differ significantly between the latter two conditions ($\chi^2(1) = 0.26$, $p = .61$). See Table 1 for the relevant numbers and percentages.

Perceived vulnerability

Because vulnerability is considered an important element of trust, and because we reasoned that having the possibility to reject B's distribution might make one feel less vulnerable to exploitation, we asked participants 'To what extent did you feel that in a situation like this, Person A is vulnerable when opting to let B distribute the chips?' (1 = absolutely not; 7 = absolutely). A one-way ANOVA on the answers to this question yielded a significant main effect of condition ($F(2,81) = 6.49$, $p = .002$; $\eta^2 = .14$). Further

Table 1. Number of participants deciding to trust vs. not to trust, per condition; Experiment 1.

	Delta = 0	Delta = 0.5 (Traditional Trust game)	Delta = 1
Trust	22 (81%)	22 (76%)	14 (50%)
Not trust	5 (19%)	7 (24%)	14 (50%)

Note: Percentages are between brackets.

testing indicated that participants reported higher levels of vulnerability in the delta = 1 condition (i.e. the condition resembling the traditional trust game; $M = 5.46$; $SD = 1.53$) than in the delta = 0.5 ($M = 4.55$; $SD = 1.96$, $t(55) = 1.96$, $p = .055$, marginally significant; Cohen's $d = 0.52$) and delta = 0 condition ($M = 3.74$; $SD = 1.81$, $t(53) = 3.82$, $p < .001$; Cohen's $d = 1.02$) condition. The latter two conditions did not differ significantly ($t(54) = 1.61$, $p = .11$; Cohen's $d = 0.42$).

Own allocations of the 20 chips

Participants who chose not to trust decided how many of the chips they allocated to themselves. Those who chose to trust were asked how they would have allocated the chips had they chosen not to trust. We first of all analysed the number of chips allocated to oneself, while disregarding the trust decision participants had made. A one-way ANOVA revealed no effect of condition ($F(2,81) = 0.44$, $p = .65$; $\eta^2 = .01$). On average, participants allocated a bit more than half of the chips to themselves ($M = 13.52$; $SD = 5.12$).

We explored whether these allocations differed for those who had not trusted (and thus made an actual allocation of the 20 chips) and those who had trusted (and thus indicated how they would have allocated the 20 chips) by adding the trust decision as a post-hoc factor. The results of this 2(decision: trusted vs. not trusted) × 3(condition) ANOVA should, of course, be interpreted with caution because on the delta = 0 and delta = .5 conditions only a small minority opted not to trust. The analysis, however, confirmed that allocations were not affected by condition ($F(2,78) = 0.38$, $p = .69$; $\eta_p^2 = .01$), nor by the trust decision ($F(1,78) = 0.68$, $p = .41$; $\eta_p^2 = .01$).

Expectations for B's distribution of 60 chips

A decision to trust implied that player B would distribute 60 chips. Here, we first analysed what decision participants expected B would make (i.e. how many chips would B allocate to A), while disregarding whether or not participants actually chose to trust or not. A one-way ANOVA on the number of chips participants thought B would allocate to them yielded a significant effect of condition ($F(2,81) = 6.66$, $p = .002$; $\eta^2 = .14$). Further testing indicated that participants expected to receive lower outcomes in the delta = 1 condition ($M = 16.43$; $SD = 8.91$), resembling the traditional trust game, than in the delta = 0.5 ($M = 21.76$; $SD = 7.08$, $t(55) = -2.51$, $p = .015$; Cohen's $d = 0.66$) and delta = 0 condition ($M = 23.96$; $SD = 7.78$, $t(53) = -3.34$, $p = .002$; Cohen's $d = 0.90$) condition. The latter two conditions did not differ significantly ($t(54) = -1,11$, $p = .27$; Cohen's $d = 0.30$).

In addition, we explored whether these expected allocations differed for those who had not trusted and those who had actually trusted B, Again, we caution against over-interpretation of this 2(decision: trusted vs. not trusted) × 3(condition) ANOVA because on the delta = 0 and delta = .5 conditions only a small minority opted not to trust. The analysis, however, confirmed that expected allocations were significantly affected by condition ($F(2,78) = 3.57$, $p = .033$; $\eta_p^2 = .084$). In addition, we only observed a main effect of decision, indicating that those who decided to trust ($M = 23.41$; $SD = 7.03$) expected B to allocate more chips to them than those who decided not to trust B ($M = 14.62$; $SD = 8.36$, $F(1,78) = 14.76$, $p < .001$; $\eta_p^2 = .16$).

Discussion

The results of our first study showed that behavioural trust differed between settings where trust means handing over all decision power to another, as in the traditional trust game, and situations that allow for deterrence-based trust. Having the possibility to reject increased A's willingness to let B distribute the chips. Cast in more theoretical terms, it appears that adding a possibility to make untrustworthy behaviour costly (by including a possibility to lower B's outcomes by rejecting his/her distribution), can increase the willingness to act on deterrence-based trust.

In addition, note that those who decided not to trust B did not favour themselves in an extreme way. On average, they allocated approximately 13 chips to themselves while allocating 7 chips to B. Moreover, those who did not trust did not allocate themselves higher outcomes than those who trusted said they would have allocated to themselves. The decision not to trust should therefore not necessarily be interpreted as a sign of competitive or self-interested behaviour.

The study also revealed some other noteworthy findings. We consistently observed differences between the delta = 1 condition and the conditions where delta was either 0.5 or 0, and found that these latter conditions did not differ significantly. This suggests a possible fundamental difference between not having (delta = 1) vs. having a possibility to reject a B's distribution (delta = 0.5 or 0), which appears to be more important than the exact magnitude of one's impact. Put differently, participants in the delta = 0.5 and delta = 0 conditions may have been more willing to trust because they had a say, which may have offered them a sense of control, rather than by a consideration of by how much exactly they could reduce the distribution. Simply having a say may be more important than the objective consequences of the say. We realise, that one could also interpret the absence of a difference between both conditions as a sign that our manipulation was not particularly strong. Note, however, that our inductions were in any case very clear and successful in the sense that all but one participant correctly identified the consequences. In other words, the differing consequences of rejecting a subsequent distribution with delta = 0 or delta = 0.5 were definitely clear enough to be well noticed and understood. But when it comes to how the consequences were subjectively perceived (in terms of seriousness of the consequences), participants perceived these differences less pronounced in the sense that they considered them equally serious.

This also does not mean that participants did not care about or did not consider the potential consequences of handing over decision power. The finding that those who did not trust had lower expectations of B does suggest the importance of expected consequences. While the current data do not allow for a discussion of whether other motives besides high expectations may steer trust, or whether people fully engage in consequential reasoning (see for that matter e.g. Dunning, Anderson, Schlösser, Ehlebracht, & Fetchenhauer, 2014; Kugler, Connolly, & Kausel, 2008), the data do seem to allow for the conclusion that participants – at least to some extent – acted on their expectations. Then again, we are also open to the suggestion that the causality may run both ways in the sense that people may also self-justify their decisions by forming post-hoc expectations. For instance, those who did not trust may have justified this by saying that they did so because they had low expectations of B (see for similar reasoning on self-justification of self-interested behaviours in other domains of interdependence e.g. Messé & Sivacek, 1979).

Experiment 2

Experiment 1 showed that the introduction of a behavioural possibility for deterrence-based trust increased the willingness to trust. In the absence of this possibility, as the findings on expectations revealed, people are more sceptical regarding B's trustworthiness. Whether or not this may partly reflect self-justifications (see above), it begs the question whether these expectations are indeed justified. How would Person B distribute the tripled chips after being trusted? Would the expectations of participants in Experiment 1 become reality? Or would the actual behaviour of Persons B prove them wrong? To test this, we designed the second experiment, in which – using the same three conditions as in Experiment 1 – all participants were assigned to the position of Person B. All participants learned that Person A had shown behavioural trust so that they could distribute the tripled amount.

Method

Design and participants

The participants, 156 students (35 males, 121 females; mean age = 21.56 years; *SD* = 2.71) at Leiden University, participated voluntarily.[7] Participants were randomly assigned to one of three conditions (delta = 1, .5, 0). No participants were excluded from the analyses.[8]

Procedure

As in Experiment 1, participants were invited to the laboratory to participate in a study on decision-making, and placed in separate cubicles, each containing a computer connected to a server. This computer was used to present the information and to register the dependent measures. At the start of the instructions, the participants were informed that they were paired with one of the other participants. Members of each dyad would be referred to as Person A and B, but now participants learned that they were assigned the letter B. The instructions of the games were identical to Experiment 1, meaning that participants learned that Person A could choose between two options: (1) They could either distribute 20 chips themselves, or (2) let B (the participant) distribute 60 chips. Participants all learned that if A would choose for option (2), A could subsequently accept or reject their distribution. In the delta = 1 condition, participants learned that the distribution would stand even if A would reject. Participants in the delta = .5 condition learned that if A would reject their distribution, it would be reduced by 50%. In the delta = 0 condition, participants learned that if A would reject the distribution, both would receive nothing. Participants learned that this information about the options was also known to A.

After this explanation, participants allegedly had to wait for A's decision. After some time (i.e. 46 seconds) participant were informed on their computer screen that A had chosen for option (2), so that they could now decide on the distribution of 60 chips. After making their distribution, additional questions were posed. Most directly related to our reasoning, we asked questions to tap how important it had been for the participant the avoid rejection of the distribution. We asked participants 'To what extent did you try to avoid that A would consider your distribution unacceptable?' and 'To what extent were you concerned with the fact that A could reject your distribution should you allocate him/her too few chips?' These questions were combined to form a reliable 'avoiding rejection' scale (Cronbach's alpha = .73).

To provide a direct check of our manipulation of delta we asked three questions, similar to those posed in Experiment 1 (but now from the perspective of player B): 'How much influence did A have if he/she would disagree with the distribution?' (1 = not at all, 7 = very much), 'What would be the consequence for you if A would disagree with your distribution?' (1 = no consequences; 7 = very serious consequences), and 'What would be the consequence for A if A would disagree with your distribution?' (1 = no consequences; 7 = very serious consequences). These questions were combined to provide a reliable 'perceived consequences of rejection' scale (Cronbach's alpha = .71).

As in Experiment 1, we also checked for understanding of the main features of the setting (i.e. whether they had been Person A or B, the value of the chips, whether decisions were anonymous, what the exact consequences of rejection would be). After these measures, the study ended. Participants were thoroughly debriefed and were paid 3.50 Euros. All participants agreed to this procedure.

Results

Manipulation checks

Out of the 156, 155 participants (99%) correctly identified their own role (Person B). Out of the 156, 152 participants (97%) correctly reported that decisions were anonymous. When asked to indicate the value of the chips (3, 5, or 10 cents), all participants correctly reported that chips were 10 cents. When asked what would be the consequences should A reject their distribution (1 = no consequences; 2 = the distribution would be reduced by 50%; we would both receive zero chips). Among 156 participants, 147 (95%) participants answered this question correctly. A one-way ANOVA for the perceived consequences scale confirmed these findings by a significant effect for condition ($F(2,155) = 52.65$, $p < .001$; $\eta^2 = .41$): The consequences should A reject B's distribution were deemed less serious in the delta = 1 condition (resembling the traditional trust game; $M = 3.18$; $SD = 1.42$) than in the delta = .5 condition ($M = 4.77$; $SD = 0.87$; $t(82.51) = -6.87$, $p < .001$, Glass's $\delta = 1.83$), which in turn yielded lower ratings than the delta = 0 condition ($M = 5.40$; $t(103) = -3.33$, $p < .001$, Cohen's $d = 0.65$).

These results indicate that the participants correctly understood the main characteristics of the experiment, and that our manipulation was successful.

Allocations

A one-way ANOVA on the allocations to Person A revealed no significant effect ($F(2, 155) = 1.93$, $p = .149$; $\eta^2 = .025$). Overall, participants allocated only a little less than half of the 60 chips to Person A ($M = 26.96$; $SD = 7.03$). The cell means are depicted in Table 2.

Motivation to prevent making an unacceptable distribution

The idea that having a final say, and thus the possibility to reject unacceptable distributions, may increase A's willingness to trust presumes that Persons B would indeed be motivated to take this into account. The allocations reported above suggest no behavioural effects, but it could have influenced the participants' motivation.[9] A one-way ANOVA on our avoiding rejection scale indeed yielded a significant effect ($F(2,155) = 18.43$, $p < .001$; $\eta^2 = .19$). This motive was reported to be less important in the delta = 1 condition (resembling the traditional trust game; $M = 3.10$; $SD = 1.50$) than in the

Table 2. Mean allocations to Person A, per condition; Experiment 2.

Delta = 0	Delta = 0.5 (Traditional Trust game)	Delta = 1
27.83 (5.56)	25.43 (9.24)	27.67 (5.39)

Note: Standard deviations are between brackets.

delta = .5 condition ($M = 4.44$; $SD = 1.47$; $t(102) = -4.62$, $p < .001$; Cohen's $d = 0.91$) and the delta = 0 condition ($M = 4.88$; $t(101) = -5.69$, $p < .001$; Cohen's $d = 1.12$). These latter two conditions did not differ significantly ($t(103) = -.44$, $p = .156$; Cohen's $d = 0.28$).

Discussion

The results of Experiment 2 are important for several reasons. First of all, the results on B's motivation showed that the manipulation of A having vs. not having a say (i.e. whether or not A could reject their distribution) did affect the participants' motivation behind the distribution they made. Thus, participants indeed seemed more motivated to prevent rejection when these consequences were more severe. Interestingly, however, the actual allocations were *not* affected by the manipulation. In other words, while affecting the underlying motive, it did not affect the resulting behaviour itself. What we saw was that participants made very generous allocations, that came close to offering an equal share of the 60 chips to A. In those cases where A could reject the distribution, this could be due to strategic, self-interested, considerations. Note, however, that in the condition resembling the traditional trust game, participants could easily have allocated the bulk of the 60 chips to themselves. We did not observe this, and as we saw in Experiment 1, players A did not anticipate such benevolent behaviour either.

The fact that we did observe a strong effect on the underlying motivations, and the fact that our manipulation checks indicated that we successfully manipulated the key construct of perceived consequences, indicates that the null effect on allocations cannot be attributed to a too weak manipulation. Indeed, note that we used exactly the same inductions that so strongly affected players A in Experiment 1; not only their willingness to trust, but also the expectations they formed regarding B's behaviour.

General discussion

How does trust work in situations of interdependence, and in particular in settings that incorporate elements of bargaining? Previous research showed that people are often too sceptical in that they trust too little (Fetchenhauer & Dunning, 2010), a conclusion that is substantiated in our current studies. With all reservations regarding comparisons between experiments, and between different measures (expectations vs. actual behaviour), it is interesting to see that – in all conditions, and even by those who showed trust – the allocations Persons A (participants in Experiment 1) expected to receive from B were lower than B's actual allocations (participants in Experiment 2).

To explain this, we feel it is important to once again consider the importance of equality. As we noted in our introduction, the preference for equality is more than 'just' the result of strategic concern, that may be reinforced when it can be used for reciprocating a positive act. A strong preference for equality is also not unique to settings such as these,

however, and one does not necessarily need reciprocity to explain such positive allocations (see e.g. Thielmann & Hilbig, 2015b). To illustrate this further, the willingness to allocate outcomes equally – even when one does not need to fear other's reactions – has previously been documented in research comparing ultimatum and delta bargaining to dictator game giving (e.g. Handgraaf et al., 2008; see also Van Dijk & Vermunt, 2000). Those findings too suggested equality as a dominant concern. Moreover, the Handgraaf et al. study revealed that the positive allocations in dictator games, which prevent any form of reciprocity, are generally not anticipated by recipients. So there too, allocators have been shown far more likely to distribute outcomes equally than anticipated by those who lacked a say in its distribution. Handgraaf et al. interpreted these findings of people behaving more nicely than expected as a manifestation of the more general phenomenon of 'egocentric empathy gaps' (see also Van Boven, Dunning, & Loewenstein, 2000), denoting that people often have difficulties taking the perspective of others. Indeed, one could say that the participants in Experiment 1 were not very successful in taking the perspective of their opponent.

The current findings speak to the importance of situational trust (cf. Li, 2007, 2012), and specifically of deterrence-based trust. The impact seems especially strong when it comes to the willingness to trust others. Having a say in the final distribution provides people with a means to deter exploitation which may help them to put their trust in others. Especially in repeated interactions, like bargaining and negotiation, which allow for mutual reciprocation, such situational features may facilitate mutual trust. Extending the traditional trust game allowed us to investigate this aspect. Future research could also compare our current one-shot setting to those of iterated games in which participants make multiple trust decisions. As the literature on repeated trust games has demonstrated, interaction partners learn from the other's behaviour, such that people become more willing to trust after having faced a trustworthy opponent, and less willing after having faced an untrustworthy opponent (e.g. Van den Bos, Van Dijk, & Crone, 2012; Yu, Saleem, & Gonzales, 2014). Our paradigm may allow for a better understanding of how in the first interaction phases trust may be enhanced by means of deterrence-based trust. For future research, it may therefore also be interesting to combine both approaches to see how introducing the option for deterrence-based trust in the early phases could be an effective means to affect trust in later stages.

At this point it is relevant to relate our findings to other research on related topics. Most relevant seems a recent study by Lenton and Mosley (2011) which showed – in a trust game setting – that having a possibility to insure oneself against non-reciprocation increased the willingness to trust. Moreover, people were willing to pay for such an insurance. In this context, it would seem plausible to expect such a willingness in our studies as well. Thus, we could envisage people being prepared to pay for having a final say. Given that our players B were in fact trustworthy regardless of A's options, this could ultimately mean that people would needlessly pay to have a final say.

When interpreting our findings, and assessing the value of generalisations such as these, one should realise that the outcomes and thus risks involved in the present study were rather small. One might wonder what would happen if higher outcomes would be at stake. This is an empirical question that may be a topic for future research. On the one hand, one might expect more self-interested behaviours from players B which would then call for a more cautious approach for players A (see also

Johansson-Stenman, Mahmud, & Martinsson, 2005; Malhotra, 2004). On the other hand, it may be noted that research on ultimatum bargaining revealed how remarkably similar allocations can be for small and large stakes (e.g. Cameron, 1999; Munier & Zaharia, 2002; but see Andersen, Ertaç, Gneezy, Hoffman, & List, 2011). We would certainly welcome such studies which – when making use of our extended trust game paradigm – could further contribute to the understanding of trust and trustworthiness in situations of interdependence.

It is also appropriate to address several other features of our paradigm, which could be viewed as limitations. First of all, it is clear that in our game, we used a binary decision to trust. As noted before, trust games (including Berg et al.'s [1995] version) often present trustors with continuous decisions. The advantage of using a binary decision – in which players had to choose between distributing 20 chips themselves or letting B distribute 60 chips, was that it allowed us to present participants in Experiment 2 with a clear outcome in which they all had to distribute the same number of chips (60), and that this would then be exactly the same as the chosen trust option of players A in Experiment 1. If we would have opted for a continuous setting in Experiment 1, this would imply that we would see more variation in the allocations made in Experiment 1. Some trustors might only transfer 3 chips (implying that player B could then distribute 9 chips), others 5 (so that player 2 could distribute 15 chips), others 8 (so that players 2 could distribute 24 chips), etcetera. This would then have the advantage of a more continuous measure for player A's trust, but would have provided a more ambiguous setting for our players B. Nevertheless, it may be interesting for future research on our extended Trust game to allow for such variation.

As one of the reviewers of this article remarked, one could think of our newly designed paradigm as a setting in which players A basically have to make a decision on which game they want to play, and what role they want to have. For example, one could say that in what we referred to as the traditional trust game setting (delta = 1), players A basically have to determine whether they (a) want to be the allocator in a 20 chips dictator game, or (b) the recipient in a 60 chips dictator game. In a similar vein, one could say that in the delta = 0 condition players A had to decide whether to (a) be an allocator in a 20 chips dictator game, or (b) a recipient in a 60 chips ultimatum game. One could then be tempted to ask: Where is the trust in all this? Is, for example the latter choice still a trust decision? Or is it primarily a strategic decision?

Previous research on behaviour in ultimatum games and delta games has indeed shown that allocators in these games are strongly affected by strategic concerns, and less so by fairness concerns. For example, research has shown that allocators in ultimatum games may distribute the money equally, not necessarily because they care for fairness, but also for strategic reasons because they fear that low offers will be rejected, leaving them empty-handed (see e.g. Kagel, Kim, & Moser, 1996; Pillutla & Murnighan, 1995; Van Dijk & Vermunt, 2000). In a similar vein, research on delta games has shown that allocators are strategic in that they respond to the consequences of rejections with lower offers as delta increases (Handgraaf et al., 2008; Suleiman, 1996).

Does this imply that trust is by definition removed from the setting as soon as you incorporate an ultimatum- or delta game in the trust game structure? Our response is that it does not. First of all, one should realise that bringing a strategic consideration to the setting does not necessarily eliminate prosocial concerns (see e.g. Van Dijk, De Cremer,

& Handgraaf, 2004). In addition, one has to acknowledge an important difference between our current setup and the setup used in isolated ultimatum bargaining studies. All participants in normal ultimatum or delta games know is that they can make an offer that can be rejected by their opponent. Importantly, they do not 'owe' this possibility to their opponent. As a result, their behaviour cannot be guided by considerations of trustworthiness because they are not in a position where they were trusted in the first place. In contrast, in our delta = 0 and delta = .5 conditions, considerations of trustworthiness can still play a role because players B who are allowed to distribute 60 chips do know that the only reason for being allowed to do this is that player A decided to give them this possibility. This information is crucial and is what brings trust to the setting. Not only for player B (who may want to prove him/herself trustworthy to A) but also to A who is likely to take this aspect into account when deciding on whether or not to make him/herself vulnerable to being a recipient in a 60 chips ultimatum game.

Our secondary data support the notion that trust is not out of the window the moment you introduce an ultimatum- or delta game structure. In Experiment 1, the data on experienced vulnerability did show that even in the delta = 0.5 and delta = 0 conditions, participants did report feeling vulnerable. In Experiment 2, the data on B's self-reported motivation did reveal that the motivation to avoid rejection was more important in the delta = 0 conditions than in the delta = 1 conditions but it did not affect their decisions. Future research may further address this matter. This could, for example, be done by testing how essential it is for players A – when placing their trust in B – that B realises that the possibility to distribute 60 chips is only there because A decided to trust him/her in the first place.

A second issue concerns the use of deception. In our two experiments we studied the behavioural decisions of players A and B independently in separate experiments. Participants were led to believe they were coupled with a counterpart, but in Experiment 1 all participants were informed that they were player A, and in Experiment 2 all participants learned that they were player B. We opted for this setup because we aimed to provide an in-depth study of both roles. As a result, we were for example able to assign over 50 participants in each condition of Experiment 2 to a setting whether they all (allegedly) had been trusted, something we could not have investigated equally well without deception (after all, then you have to rely on whether or not participants indeed are willing to trust). The implication was that in the end, payments were not contingent on decisions (because participants were not actually connected to their counterpart). While such a research approach is not unusual in the field of psychology, we realise that the use of deception and noncontingent payment is considered problematic in the fields of experimental economics (see for more elaborate discussions of the differences between fields e.g. Ariely & Norton, 2007; Croson, 2005; Hertwig & Ortmann, 2001). For future research, and further integration of the fields, it seems worthwhile to test our ideas in non-deception contexts (e.g. by matching players B to real players A) and with contingent rewards. As Johnson and Mislin's (2011) meta-analysis suggests, absolute levels of trust may be affected by such inductions. The current literature does not yet allow for a prediction of whether these inductions would have a different effect for trust settings that allow vs. not allow for deterrence-based trust (e.g. whether in our settings it would have a different impact in the delta = 0 vs. delta = 1 conditions), but this would definitely be an interesting question to pursue.

As a final note, we wish to point to yet another interesting avenue for future research. In a recent study, Brion, Lount, and Doyle (2015) introduced a new issue to the field of trust research by raising the question of how meta-perceptions about trustworthiness affect trust development. How trustworthy do others think I am? Being aware of how others view one's trustworthiness seems essential to the development of trust. In the current paper we showed how in a bargaining setting having an option to reject distributions can influence expectations of trustworthiness without affecting actual trustworthiness. An interesting next step would then be to see whether meta-perceptions would change as well: Do people anticipate the effects that options to accept/reject can have on their perceived (deterrence-based) trustworthiness? Given that accuracy may facilitate the development and building of trust this may very well prove to be another meaningful contribution to the study and understanding of trust.

Notes

1. Note that in this respect, the setup also slightly differs from Berg et al.'s (1995) study, where a no-trust decision implies that all money will be retained by Person A. In our setup, A makes an additional decision on how to distribute these outcomes.
2. See note 1 above.
3. Exclusion of participants who did not correctly answer one of the checks did not change the findings: All significant effects we report remained significant, and all non-significant effects remained non-significant.
4. We report the results on the dependent measures that are most central to our reasoning; the checks on the manipulations, main dependent measures, and main motives. For exploratory reasons we also collected additional data, for example, on social value orientations and reciprocity orientation (both taken prior to the experiments) and other data assessed during the experiment (e.g. assessments of power). These – more remotely related – data are available upon request. In both studies, the trust decision was the first measure taken in the experiment (i.e. in Experiment 1, the decision to trust or not; in Experiment 2 the first measure was the participants' distribution of the 60 chips).
5. See note 3.
6. See note 4.
7. Prior to running Experiment 2, we also ran the same experiment with 64 participants. The results of this lower powered study were identical to the findings we now report, lending more confidence in the robustness of our findings. As in Experiment 2 we report here, the data showed the success of our manipulations on then manipulation checks, no effect of our manipulations on allocations, and an effect on the reported motivation to avoid rejection, with the mean levels being practically identical. These data are available upon request. The number of participants for the current Experiment 2 was based on A power analysis (G*Power 3.1; Faul, Erdfelder, Buchner, & Lang, 2009) indicated that 159 participants would yield a statistical power of .80 (with an alpha of 0.05 an a medium sized effect, $f = .25$; see Cohen, 1988).
8. For three participants, the age data were not correctly stored by the computer program, so the mean age is based on 153 participants. Exclusion of participants who did not correctly answer one of the manipulations checks did not change the findings: All significant effects we report remained significant, and all non-significant effects remained non-significant.
9. We observed a small but non-significant positive correlation between both measures ($r = .11$, $p = .19$).

Acknowledgements

We thank Carlo van Loenhout for his contributions to the project.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Eric van Dijk is professor at the Department of Social and Organizational Psychology of Leiden University. His research interests include uncertainty, risk taking, emotions, bargaining, and social dilemmas. He has served as the Scientific Director of the Kurt Lewin Institute, and was associate editor for the *British Journal of Social Psychology* and *Organizational Behavior and Human Decision Processes*.

Varia Makagonova was graduated from the Master specialization Economic and Consumer Psychology at Leiden University. She currently works as Global Communications Manager at Hippo, a BloomReach company.

Erik W. de Kwaadsteniet is assistant professor at the Department of Social and Organizational Psychology of Leiden University. His research interests include social dilemmas, sanctions, and tacit coordination.

Manon Schutter was graduated from the Master specialization Economic and Consumer Psychology at Leiden University, and is currently Ph.D. student on a project to study the dynamics of trust.

References

Andersen, S., Ertaç, S., Gneezy, U., Hoffman, M., & List, J. A. (2011). Stakes matter in ultimatum games. *American Economic Review*, *101*, 3427–3439. doi:10.1257/aer.101.7.3427

Ariely, D., & Norton, M. I. (2007). Psychology and experimental economics: A gap in abstraction. *Current Directions in Psychological Science*, *16*, 336–339. doi:10.1111/j.1467-8721.2007.00531.x

Ashraf, N., Bohnet, I., & Piankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, *9*, 193–208. doi:10.1007/s10683-006-9122-4

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*, 122–142. doi:10.1006/game.1995.1027

Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, *55*, 467–484. doi:10.1016/j.jebo.2003.11.004

Bolton, G. E., Katok, E., & Zwick, R. (1998). Dictator game giving: Rules of fairness versus acts of kindness. *International Journal of Game Theory*, *27*, 269–299. doi:10.1007/s001820050072

Brion, S., Lount, Jr, R. B., & Doyle, S. P. (2015). Knowing if you are trusted: Does meta-accuracy promote trust development? *Social Psychological and Personality Science*, *6*, 823–830. doi:10.1177/1948550615590200

Cameron, L. A. (1999). Raising the stakes in the ultimatum game: Experimental evidence in Indonesia. *Economic Inquiry*, *37*, 47–59. doi:10.1111/j.1465-7295.1999.tb01415.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Croson, R. (2005). The method of experimental economics. *International Negotiation*, *10*, 131–148. doi:10.1163/1571806054741100

Dunning, D., Anderson, J. A., Schlösser, T., Ehlebracht, D., & Fetchenhauer, D. (2014). Trust at zero acquaintance: More a matter of respect than expectation of reward. *Journal of Personality and Social Psychology*, *107*, 122–141. doi:10.1037/a0036673

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. doi:10.3758/BRM.41.4.1149

Fetchenhauer, D., & Dunning, D. (2009). Do people trust too much or too little? *Journal of Economic Psychology*, *30*, 263–276. doi:10.1016/j.joep.2008.04.006

Fetchenhauer, D., & Dunning, D. (2010). Why so cynical? Asymmetric feedback underlies misguided skepticism regarding the trustworthiness of others. *Psychological Science*, *21*, 189–193. doi:10.1177/0956797609358586

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, *115*. doi:10.1162/003355300554926

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, *3*, 367–388. doi:10.1016/0167-2681(82)90011-7

Handgraaf, M. J. J., Van Dijk, E., Vermunt, R. C., Wilke, H. A. M., & De Dreu, C. K. W. (2008). Less power or powerless? Egocentric empathy gaps and the irony of having little versus no power in social decision making. *Journal of Personality and Social Psychology*, *95*, 1136–1149. doi:10.1037/0022-3514.95.5.1136

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences*, *24*, 383–403.

Johansson-Stenman, O., Mahmud, M., & Martinsson, P. (2005). Does stake size matter in trust games? *Economics Letters*, *88*, 365–369. doi:10.1016/j.econlet.2005.03.007

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, *32*, 865–889. doi:10.1016/j.joep.2011.05.007

Kagel, J. H., Kim, C., & Moser, D. (1996). Fairness in ultimatum games with asymmetric information and asymmetric payoffs. *Games and Economic Behavior*, *13*, 100–110. doi:10.1006/game.1996.0026

Kreps, D. M. (1990). Corporate culture and economic theory. In J. Alt & K. Shepsle (Eds.), *Perspectives on positive political economy* (pp. 90–143). Cambridge: Cambridge University Press.

Kugler, T., Connolly, T., & Kausel, E. E. (2008). The effect of consequential thinking on trust game behavior. *Journal of Behavioral Decision Making*, *22*, 101–119. doi:10.1002/bdm.614

Lenton, P., & Mosley, P. (2011). Incentivising trust. *Journal of Economic Psychology*, *32*, 890–897. doi:10.1016/j.joep.2011.07.005

Lewicki, R. J., & Bunker, B. B. (1996). Developing and maintaining trust in work relationships. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 133–174). Thousand Oaks, CA: Sage Publications.

Lewicki, R. J., & Polin, B. (2013). The role of trust in negotiation processes. In R. Bachmann & A. Zaheer (Eds.), *Handbook of advances in trust research* (pp. 29–54). Cheltenham: Edward Elgar.

Lewicki, R. J., Tomlinson, E. C., & Gillespie, N. (2006). Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *Journal of Management*, *32*, 991–1022. doi:10.1177/01492063294405

Li, P. P. (2007). Toward and interdisciplinary conceptualization of trust: A typological approach. *Management and Organization Review*, *3*, 421–445.

Li, P. P. (2012). When trust matters the most: The imperatives for contextualising trust research. *Journal of Trust Research*, *2*, 101–106. doi:10.1080/21515581.2012.708494

Malhotra, D. (2004). Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties. *Organizational Behavior and Human Decision Processes*, *94*, 61–73. doi:10.1016/j.obhdp.2004.03.001

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*, 709–734. doi:10.2307/258792

Messé, L. A., & Sivacek, J. M. (1979). Predictions of others' responses in a mixed-motive game: Self-justification or false consensus? *Journal of Personality and Social Psychology*, *37*, 602–607. doi:10.1037//0022-3514.37.4.602

Messick, D. M. (1995). Equality, fairness, and social conflict. *Social Justice Research*, *8*, 153–173.

Munier, B., & Zaharia, C. (2002). High stakes and acceptance in ultimatum bargaining: A contribution from an international experiment. *Theory and Decision*, *53*, 187–207. doi:10.1023/A:1022815832351

Olekalns, K., Kulik, C. T., & Chew, L. (2014). Sweet little lies: Social context and the use of deception in negotiation. *Journal of Business Ethics*, *120*, 13–26. doi:10.1007/s10551-013-1645-y

Olekalns, M., Lau, F., & Smith, P. L. (2007). Resolving the empty core: Trust as a determinant of outcomes in three-party negotiations. *Group Decision and Negotiation*, *16*, 527–538. doi:10.1007/s10726-007-9084-8

Pillutla, M. M., & Murnighan, J. K. (1995). Being fair or appearing fair: Strategic behavior in ultimatum game bargaining. *Academy of Management Journal*, *38*, 1408–1426. doi:10.2307/256863

Rousseau, D., Sitkin, S., Burt, R., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*, 393–404. doi:10.5465/AMR.1998.926617

Schilke, O., Reimann, M., & Cook, K. S. (2015). Power decreases trust in social exchange. *Proceedings of the National Academy of Sciences*, *112*, 12950–12955. doi:10.1073/pnas.1517057112

Shapiro, D., Sheppard, B. H., & Cheraskin, L. (1992). Business on a handshake. *Negotiation Journal*, *8*, 365–377. doi:10.1111/j.1571-9979.1992.tb00679.x

Suleiman, R. (1996). Expectations and fairness in a modified ultimatum game. *Journal of Economic Psychology*, *17*, 531–554. doi:10.1016/S0167-4870(96)00029-3

Thielmann, I., & Hilbig, B. E. (2015a). Trust: An integrative review from a person-situation perspective. *Review of General Psychology*, *19*, 249–277. doi:10.1037/gpr0000046

Thielmann, I., & Hilbig, B. E. (2015b). The traits one can trust: Dissecting reciprocity and kindness as determinants of trustworthy behavior. *Personality and Social Psychology Bulletin*, *41*, 1523–1536. doi:10.1177/0146167215600530

Van Boven, L., Dunning, D., & Loewenstein, G. (2000). Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. *Journal of Personality and Social Psychology*, *79*, 66–76. doi:10.1037//0022-3514.79.1.66

Van den Bos, W., Van Dijk, E., & Crone, E. C. (2012). Learning whom to trust in repeated social interactions: A developmental perspective. *Group Processes & Intergroup Relations*, *15*, 243–256. doi:10.1177/1368430211418698

Van Dijk, E., De Cremer, D., & Handgraaf, M. J. J. (2004). Social value orientations and the strategic use of fairness in ultimatum bargaining. *Journal of Experimental Social Psychology*, *40*, 697–707. doi:10.1016/j.jesp.2004.03.002

Van Dijk, E., & Vermunt, R. (2000). Strategy and fairness in social decision making: Sometimes it pays to be powerless. *Journal of Experimental Social Psychology*, *36*, 1–25. doi:10.1006/jesp.1999.1392

Van Dijk, E., Wit, A. P., Wilke, H., & De Kwaadsteniet, E. W. (2010). On the importance of equality in social dilemmas. In R. M. Kramer, M. H. Bazerman, & A. E. Tenbrunsel (Eds.), *Social decision making: Social dilemmas, social values, and ethical judgments* (pp. 47–69). New York: Routledge.

Yu, M., Saleem, M., & Gonzales, C. (2014). Developing trust: First impressions and experience. *Journal of Economic Psychology*, *43*, 16–29. doi:10.1016/j.joep.2014.04.004