Can the Trust Game Paradigm Really Measure Trust? Controversies and Prospects

Authors: Gong Zhe, Tang Yujie, Liu Chang, Liu Chang

Date: 2020-08-29T00:00:00+00:00

Abstract

Trust is a psychological state based on positive expectations of others' intentions or behaviors, where one is willing to entrust (i.e., willing to bear risks) (张宁等人, 2011). For a long time, trust measurement has been a central and challenging issue in trust research, and researchers have failed to reach a consensus on how to properly and accurately measure trust (Brülhart & Usunier, 2012; Corgnet et al., 2016; Gylfason & Olafsdottir, 2017).

According to Bauer and Freitag (2018), trust measurement is divided into direct measurement and indirect measurement. Direct measurement refers to participants directly reporting their trust in a certain dimension on a Likert scale, such as the Interpersonal Trust Scales (ITS) (Rotter, 1967); indirect measurement infers an individual's trust level through cognitive decision-making, specific behaviors, etc. Among indirect measurement methods for trust, the most prominent is the trust game paradigm (trust game) developed by Berg et al. (1995) based on Deutsch' s (1958) prisoner' s dilemma game, hereinafter referred to as the trust game. The trust game includes two roles: trustor and trustee. In the classic trust game, both parties are endowed with a certain amount of real or virtual currency, the trustor can invest any amount of money to the trustee, and the trustee receives a multiplied gain (typically two or three times the investment); subsequently, the trustee can return any amount of money from the gained amount to the trustor. In this game context, the amount invested by the trustor to the trustee (hereinafter collectively referred to as investment trust level) represents trust, and the trustee' s return amount represents the trustee' s trustworthiness.

The trust game has been widely used (Espín et al., 2016; Bellucci et al., 2019; Javor et al., 2015), but its internal validity and ecological validity have been subject to considerable scrutiny (Karlan, 2005; Thielmann & Hilbig, 2015). Some researchers even doubt that investment behavior in the trust game can measure

trust at all (Ermisch & Gambetta, 2006; Sapienza et al., 2013). Given the important role of the trust game in trust research, it is of great significance to review the main controversies related to it, analyze and evaluate the reasonableness of these controversies, and propose prospects for future research.

Currently, the main controversies facing the trust game include two aspects. First, since its establishment by Berg et al. (1995), the trust game paradigm has undergone numerous variations, but the reasonableness of these paradigm variations remains questionable (Aksoy et al., 2018; Lenton & Mosley, 2011). Research has found that investment trust level may be influenced by social preference and risk preference during paradigm variations (Aksoy et al., 2018; Derks et al., 2014; Evans & Krueger, 2011; Thielmann & Hilbig, 2015), thereby reducing the internal validity of the trust game. Second, the correlation between investment trust level in the trust game and survey-based trust (hereinafter collectively referred to as survey trust) measured by scale methods is relatively low (Ashraf et al., 2006; Etang et al., 2012; Holm & Nystedt, 2008). Some researchers have thus questioned that the validity of the trust game is not reliable (Brülhart & Usunier, 2012; Karlan, 2005). The following sections will first organize and analyze the main controversies of the trust game from these two perspectives, and finally propose prospects for future research based on this analysis.

Full Text

Preamble

Can the Trust Game Paradigm Really Measure Trust? Controversies and Prospects

GONG Zhe[1], TANG Yujie[1], LIU Chang[2]
([1] School of Psychology, Sichuan Normal University, Chengdu 610068, China)
([2] School of Psychology, Nanjing Normal University, Nanjing 210097, China)

Abstract

The validity of the trust game paradigm for measuring trust has long been controversial. On one hand, social preferences and risk preferences may influence investment trust levels during paradigm variations, reducing the internal validity of the trust game. On the other hand, the correlation between trust game performance and survey-based trust is relatively low, which may be related to differences between attitudes and behaviors, measurement type differences, and limitations inherent in the measurement paradigms themselves. Despite these controversies, the trust game paradigm remains a relatively appropriate method for trust measurement. Future improvements to the trust game paradigm should focus on: verifying the scientific validity of paradigm variations; further clarifying the reasons for low correlations with survey trust; expanding the measurement dimensions of the trust game paradigm; and enhancing its ecological

validity.

Keywords: trust; trust game; social preference; risk preference

1 Introduction

Trust is a psychological state of willingness to take risks by entrusting others based on positive expectations of their intentions or behaviors (Zhang et al., 2011). For a long time, trust measurement has been a key and challenging issue in trust research, with researchers struggling to reach consensus on how to properly and accurately measure trust (Brülhart & Usunier, 2012; Corgnet et al., 2016; Gylfason & Olafsdottir, 2017). According to Bauer and Freitag (2018), trust measurement can be divided into direct and indirect approaches. Direct measurement requires participants to report their trust levels on a Likert scale for specific dimensions, such as the Interpersonal Trust Scales (ITS) (Rotter, 1967). Indirect measurement infers trust levels through individuals' cognitive decisions and specific behaviors. Among indirect trust measurement methods, the most prominent is the trust game paradigm developed by Berg et al. (1995) based on Deutsch' s (1958) prisoner' s dilemma game, hereinafter referred to as the trust game. The trust game involves two roles: the trustor and the trustee. In the classic trust game, both parties are endowed with a certain amount of real or virtual currency. The trustor can invest any amount with the trustee, who receives a multiplied value (typically two or three times the investment). Subsequently, the trustee can return any portion of this multiplied amount to the trustor. In this game scenario, the amount invested by the trustor (hereinafter referred to as investment trust level) represents trust, while the amount returned by the trustee represents trustworthiness.

Although the trust game has been widely used (Espín et al., 2016; Bellucci et al., 2019; Javor et al., 2015), its internal and ecological validity has faced numerous challenges (Karlan, 2005; Thielmann & Hilbig, 2015). Some researchers have even questioned whether investment behavior in the trust game can measure trust at all (Ermisch & Gambetta, 2006; Sapienza et al., 2013). Given the important role of the trust game in trust research, reviewing its main controversies, analyzing and evaluating their validity, and proposing future research directions are of significant importance.

Currently, the main controversies surrounding the trust game include two aspects. First, since its establishment by Berg et al. (1995), the trust game paradigm has undergone many variations, yet the rationality of these variations remains questionable (Aksoy et al., 2018; Lenton & Mosley, 2011). Research has found that investment trust levels may be influenced by social preferences and risk preferences during paradigm variations (Aksoy et al., 2018; Derks et al., 2014; Evans & Krueger, 2011; Thielmann & Hilbig, 2015), thereby reducing the internal validity of the trust game. Second, the correlation between investment trust levels in the trust game and survey-based trust (hereinafter referred to as survey trust) is relatively low (Ashraf et al., 2006; Etang et al., 2012; Holm

& Nystedt, 2008). Some researchers have questioned the reliability of the trust game' s validity based on this low correlation (Brülhart & Usunier, 2012; Karlan, 2005). The following sections will first analyze the main controversies surrounding the trust game from these two perspectives, and then propose future research directions based on this analysis.

2 Variations in the Trust Game Paradigm

Based on Berg et al.' s (1995) classic paradigm, researchers have mainly made variations in four aspects: the equality of initial endowments, the continuity of investment amounts, the amount available for investment, and the investment multiplication rate (Carragher et al., 2018; Johnson & Mislin, 2011; Woolley & Fishbach, 2017).

2.1 Equality of Initial Endowments

The classic trust game provides equal initial amounts to both trustors and trustees. However, some researchers, considering cost constraints, have eliminated the trustee' s initial endowment (Schlösser et al., 2015; Woolley & Fishbach, 2017), creating inequality in initial payoffs. This inequality may induce feelings of guilt or distress in trustors, leading them to invest more based on inequity aversion rather than trust intentions (Aksoy et al., 2018). Conversely, Evans and Revelle (2008) noted that unequal initial endowments might also suppress trustors' investment levels, as the inequality relatively enhances the trustor' s investment capacity, potentially leading them to reduce investment trust levels for self-interested reasons. A meta-analysis by Johnson and Mislin (2011) encompassing 162 studies supported Evans and Revelle' s (2008) speculation, finding that trustors invested significantly less when initial payoffs were unequal compared to when they were equal.

Furthermore, the effect of initial endowment equality on investment trust levels may be moderated by trustor characteristics. For example, Corgnet et al. (2016) found that individuals who more strongly disliked unequal initial endowments exhibited higher investment trust levels. Additionally, Hong and Bohnet (2007) showed that unequal initial endowments only suppressed investment trust levels among lower-status groups in the U.S. population—women, ethnic minorities, young people, and non-Protestants. This suggests that trustors may differentially interpret unequal initial endowments based on factors such as social status and fairness beliefs, resulting in varying investment trust levels.

2.2 Continuity of Investment Amounts

The classic trust game uses continuous investment amounts (Castilla, 2015; Gereke et al., 2018), allowing participants to choose any number within the investment range. While widely used, this continuous setting has revealed drawbacks. Research shows that participants tend to give "moderate" investments that are neither too high nor too low (Bellucci et al., 2019; McAuliffe et al.,

2019). This tendency may be unrelated to trust and instead reflect altruistic or fairness motivations (Ashraf et al., 2006; Rudnicki et al., 2019). To circumvent this issue, some researchers have developed binary trust game paradigms that force a choice between "invest all" or "invest nothing" (Carragher et al., 2018; Evans & Krueger, 2011).

Evans and Krueger' s (2011) sequential trust game paradigm exemplifies this approach. The paradigm includes trustor and trustee roles with four payoff levels ordered as $T > R > P > S$. After the game begins, the trustor must choose between trusting or maintaining the status quo. Under the status quo condition, the game ends immediately, with both players receiving the lower payoff P. Under the trust condition, the game proceeds to the second stage, where the trustee chooses between reciprocity and betrayal. If the trustee chooses reciprocity, both players receive the higher payoff R. If betrayal is chosen, the trustor receives the lowest payoff S while the trustee receives the highest payoff T. According to the game rules and payoff matrix, the trustor' s decision to trust may yield higher returns than maintaining the status quo ($R > P$) but also risks lower returns due to betrayal ($S < P$). Therefore, the trustor must evaluate whether the risk of betrayal is worth taking before deciding to trust or maintain the status quo. Researchers speculate that binary trust game paradigms can prevent investment trust behavior from being overly interpreted as social preferences (Carragher et al., 2018; Yamagishi et al., 2015).

2.3 Amount Available for Investment and Investment Multiplication Rate

Researchers differ in the amounts they provide participants for investment. Most Western studies set amounts between $5 and $15 (Bellucci et al., 2019; Javor et al., 2015; Müller & Schwieren, 2019), with some below $4 (Woolley & Fishbach, 2017) or above $20 (Drążkowski et al., 2017). Chinese studies typically use amounts between ¥5 and ¥20 (Dou et al., 2018; Fu et al., 2018; Liu et al., 2017). We argue that the setting of available investment amounts is related to regional characteristics, economic features, and other external environmental factors in the research context, and that regional differences are reasonable. The problem lies in the arbitrary nature of amount settings in current research, lacking proper justification. This may cause investment trust levels to be significantly influenced by the amount setting, thereby reducing the internal validity of the trust game. Research shows that when trustors have more available funds, their potential losses from investment increase, leading to lower investment amounts (Evans & Krueger, 2011). Johansson-Stenman et al. (2013) also demonstrated that investment trust levels significantly decline when available investment amounts increase. Johnson and Mislin' s (2011) meta-analysis similarly found that the amount available for investment significantly predicts investment trust levels.

The investment multiplication rate refers to the rate at which the trustor' s investment is multiplied. Berg et al.' s (1995) classic paradigm uses a threefold

multiplication. Some researchers have changed this to twofold (Gereke et al., 2018; Woolley & Fishbach, 2017) or set it to a specific amount below threefold, such as multiplying $2 into $5 (Levine & Schweitzer, 2015). The investment multiplication rate is closely related to investment trust levels. When the multiplication rate increases, the theoretical return on investment rises, making participants more willing to take risks and exhibit investment trust behavior. Conversely, participants' investment trust motivation may be suppressed (Evans & Krueger, 2011; Johnson & Mislin, 2011). For example, Lenton and Mosley (2011) examined the relationship between investment multiplication rate and investment trust levels, finding that investment trust levels significantly increased with higher multiplication rates. In scenarios with twofold, threefold, and fourfold multiplication rates, participants' investment amounts as a proportion of initial endowments were 30%, 48%, and 53%, respectively.

2.4 Impact of Paradigm Variations on Measurement Validity

Variations in the trust game paradigm provide new avenues for expanding its application and enhancing research validity (e.g., binary trust games). However, some variations lack sufficient empirical support and retain a degree of arbitrariness. For instance, researchers' decisions to reduce available investment amounts or eliminate trustees' initial endowments, while saving experimental costs, have sparked considerable controversy. The core of such controversy concerns the relationship between paradigm variations and measurement validity. As a trust measurement paradigm, trust behavior in the trust game should be driven by trust intentions. However, when paradigm variations occur, trustors' investment behavior may be more strongly explained by other variables (e.g., social preferences, risk preferences), reducing the internal validity of the trust game (Bicchieri et al., 2011). Therefore, researchers'controversies over paradigm variations essentially reflect concerns about how social preferences and risk preferences influence investment trust levels.

Numerous studies have found that investment trust levels are closely related to social preferences (Aksoy et al., 2018; Derks et al., 2014; Johansson-Stenman et al., 2013; Sapienza et al., 2013) and risk preferences (Fetchenhauer & Dunning, 2012; Thielmann & Hilbig, 2015), leading some researchers to question whether the trust game can adequately represent trust (Ermisch & Gambetta, 2006; Sapienza et al., 2013). The following sections analyze the impact of social preferences and risk preferences on the trust game.

2.4.1 Influence of Social Preferences Social preferences are primarily related to controversies over investment amount equality and continuity settings. Based on the view that individual cooperativeness influences trust behavior (e.g., through social projection), social preferences represented by altruism, reciprocity, and inequity aversion are considered important predictors of trust (Aksoy et al., 2018; Derks et al., 2014; Johansson-Stenman et al., 2013). For example, Yamagishi et al. (2015) found that prosocial preferences were significantly

positively correlated with investment trust levels. Naef and Schupp (2009) showed that participants who more frequently engaged in volunteer services exhibited higher investment trust levels.

Regarding the mechanism through which social preferences influence investment trust, strong reciprocity theory posits that people derive utility satisfaction and benefits from the process of reciprocal behavior itself (Gintis, 2000). Some researchers argue that in the trust game, the act of investing in others includes a psychological process of "giving" to others, from which individuals gain a sense of efficacy (Ozyilmaz et al., 2018). In Dunn et al.' s (2008) classic study published in *Science*, both survey data and experimental results showed that people feel happier spending money on others than on themselves. This suggests that investment behavior itself is influenced by prosocial and inequity aversion factors. Therefore, using the act of "giving" money to others in the trust game to represent trust may confound trust and social preferences (Espín et al., 2016).

To clarify this issue, researchers have conducted a series of validation studies. Although the close relationship between social preferences and investment trust has been frequently observed (Fairley et al., 2016; Rudnicki et al., 2019), many studies do not support the hypothesis that social preferences significantly influence investment trust behavior (Ashraf et al., 2006; Bellucci et al., 2019; Brülhart & Usunier, 2012; Javor et al., 2015). For example, in Brülhart and Usunier's (2012) study, dominant altruism theory would predict that as trustors, participants should invest more in trustees with smaller initial endowments and less in those with larger endowments. However, results showed no significant correlation between investment amounts and trustees' initial endowments. Overall, the relationship between social preferences and investment trust levels remains unclear and requires further investigation. Nevertheless, given the potential role of social preferences, researchers often measure them in experiments to assess their impact on results through statistical analysis (Corgnet et al., 2016; Schlösser et al., 2015).

2.4.2 Influence of Risk Preferences Risk preferences are primarily related to controversies over available investment amounts and multiplication rate settings. In economics, risk preference refers to individuals' psychological reactions, attitudinal tendencies, and investment willingness when facing risk during investment choices (Markowitz, 1952), comprising a three-dimensional structure of risk aversion, risk proneness, and risk neutrality (Pratt, 1964). From the perspective of traditional economic game theory, individuals' behavior in the trust game resembles venture investment (Fehr, 2009) and is susceptible to risk aversion or risk-seeking factors (Evans & Krueger, 2011; Fairley et al., 2016; Thielmann & Hilbig, 2015). Therefore, the amount individuals invest in others in the trust game may represent their risk preference levels (Castilla, 2015; Müller & Schwieren, 2019).

On one hand, research shows that individuals' general attitudes toward uncertainty (i.e., risk aversion) and potential losses (i.e., loss aversion) influence their

willingness to trust others (Karlan, 2005; Thielmann & Hilbig, 2015). Sapienza et al. (2013) found that a one-standard-deviation decrease in risk aversion increased average investment trust amounts by approximately 17%. Similar findings indicate that individuals' aversion to betrayal in venture investments leads them to prefer lower risks when investment returns are determined by others rather than randomly (Fairley et al., 2016). On the other hand, individuals' risk-seeking tendencies drive them to pursue benefits and exhibit investment trust behavior. For example, Chetty et al. (in press) found that individuals who preferred risk-seeking in lottery choice tasks invested more in the trust game.

However, many studies have not found significant effects of risk preferences on investment trust levels (Ben-Ner & Halldorsson, 2010; Fetchenhauer & Dunning, 2012). For instance, Ben-Ner and Halldorsson (2010) showed that neither survey-based nor behavioral measures of risk aversion were significantly correlated with investment trust levels. Dunning et al. (2014) found that based on self-reported risk tolerance, only 20.4% of participants were willing to invest in others in the trust game, yet 62.4% actually invested. This suggests that the influence of risk preferences on the trust game is conditional. Considering the potential impact of risk preferences on investment trust levels, many researchers measure risk preferences in experiments to assess their influence on results (Breuer et al., 2016; Yamagishi et al., 2015).

2.5 Summary

Overall, social preferences and risk preferences have non-universal effects on investment trust levels, and the degree of influence is associated with variations in the trust game paradigm. However, the impact of paradigm variations on research validity and the roles of social preferences and risk preferences remain unclear and require further investigation.

Although controversies over the influence of social preferences and risk preferences have raised questions about the trust game (Aksoy et al., 2018; Rudnicki et al., 2019), the close relationship between investment trust levels and social/risk preferences is theoretically reasonable given trust' s multidisciplinary and multi-perspective research status. From a conceptual perspective, some researchers view trust as essentially a moral concept or basic norm in social contexts (Bicchieri et al., 2011; Uslaner, 2002; Reiersen, 2019). For example, Uslaner (2002) proposed that trust is a stable value—a moral imperative assuming others are trustworthy—that does not easily change due to specific events or benefit calculations. Fukuyama (2001) also argued that trust in society represents a group' s universal recognition of a value system. Accordingly, trust motivation includes social preferences such as altruism, making investment trust levels highly susceptible to social preference influences. On the other hand, according to the "economic man" hypothesis, trust behavior is essentially an investment behavior in risky situations, making it inevitably related to risk preferences. From a game theory perspective, traditional game theory suggests that both parties in a prisoner' s dilemma can achieve "Nash equilibrium" through mutual distrust

(Ma et al., 2009). However, individuals in trust games tend to express mutual trust, deviating from profit-maximizing behavior. This may imply that the pursuit of material wealth is not humanity' s only motivation; human behavior is also driven by intrinsic and extrinsic motivations that value morality, emotions, and social norms (Jiang & Wei, 2013). This behavioral game theory approach can lead to new "Nash equilibrium solutions" to explain many paradoxes found in behavioral game experiments (Ma et al., 2009). Therefore, the close relationship between investment trust levels and social/risk preferences also deepens our understanding of the trust concept. In summary, social preferences and risk preferences may constitute sources of trust intentions underlying investment trust behavior, and questioning the validity of the trust game based solely on their positive relationships with investment trust levels is biased. However, this does not mean the trust game is equivalent to risk preference or social preference paradigms; it remains essentially a trust measurement paradigm. The key is to develop and apply the trust game paradigm reasonably, avoiding direct interpretation of investment trust behavior as social or risk preferences, thereby improving internal validity.

3 The Low Correlation with Survey Trust

Some researchers administer survey trust measures alongside the trust game to provide criterion-related validity evidence (Etang et al., 2012; Yamagishi et al., 2015). However, many studies show that survey trust lacks stable, reliable correlations with investment trust (Ashraf et al., 2006; Holm & Nystedt, 2008). Meanwhile, participants' survey trust levels have been found to correlate more significantly with the amount they return when playing as trustees—that is, with trustworthiness (Karlan, 2005; McEvily et al., 2012). For example, Karlan (2005) found that individuals' survey trust levels significantly predicted their trustworthiness but not their investment trust performance. The low correlation phenomenon between the trust game and survey trust has attracted continuous attention, with some researchers arguing that the trust game lacks validity (Brülhart & Usunier, 2012; Karlan, 2005). We argue that this low correlation can be explained from three perspectives: attitude-behavior differences, measurement type differences, and limitations of both survey trust and the trust game.

3.1 Attitude-Behavior Differences

Regarding attitude-behavior differences, survey trust mostly uses self-report methods requiring participants to rate statements on Likert scales, reflecting individuals' attitudinal tendencies. Therefore, survey trust is also interpreted as trust attitude (Gächter et al., 2004). In contrast, the trust game represents trust behavior. While attitudes and behaviors are closely related, their relationship is not deterministic (Shirokova et al., 2016). Some researchers even argue that the direct correlation between attitudes and behaviors is so low that it poses an unprecedented challenge to psychologists (Cooper & Croyle, 1984).

The low correlation between survey trust and the trust game can be discussed within the framework of attitude-behavior relationships.

Research shows that individuals' intentions, attitude strength, and interpersonal relationships can all lead to inconsistencies between attitudes and behaviors (Smith & Louis, 2008). The study of attitude-behavior relationships has evolved from the Theory of Reasoned Action (TRA) to the Theory of Planned Behavior (TPB), which incorporates irrational factors (Michaelidou & Hassan, 2014). Researchers have gradually recognized that attitudes are complex structures comprising both affect and cognition, while behaviors are not entirely driven by rationality. Behaviors only reflect rational beliefs, attitudes, and intentions, and in some special cases, cannot be measured by rational standards (Zhang & Wang, 2007). For example, research has found that implicit and explicit attitudes relate differently to behaviors: implicit attitudes are more closely related to fast, automatic behaviors, while explicit attitudes correlate more strongly with consciously controlled behaviors (Swanson et al., 2001). This reveals that the complex relationship between behavioral and attitudinal systems is influenced by both rational and irrational variables. Similarly, trust has evolved from a purely rational view to an irrational view and then to a bounded rationality view (Xu & Liu, 2007). Trust behavior is related to a series of factors including trustor characteristics, trustee characteristics, and situational variables that encompass both rational and irrational elements, and its relationship with trust attitude is influenced by numerous factors (Bujang et al., 2017). Therefore, the low correlation between survey trust attitudes and trust game behaviors may result from the uncertain relationship between the attitudinal and behavioral systems they represent.

3.2 Measurement Type Differences

Some researchers argue that the lack of stable correlation between survey trust and the trust game may stem from inconsistent trust types and trust objects represented by the two measurement methods in specific studies (Sapienza et al., 2013). From a domain general-specific perspective, trust can be divided into general trust and specific trust. General trust typically represents individuals' universal, relatively stable, and objectless trust propensity or trait. For example, the classic question: "Generally speaking, would you say that most people can be trusted or that you can' t be too careful in dealing with people?" appears in trust surveys such as the World Values Survey and General Social Survey (GSS). In contrast, specific trust generally refers to trust in particular or specific trust objects, such as trust in doctors (Kovacs et al., 2019) or parents (Kerr et al., 1999).

Researchers often use survey trust to represent general trust (Bibi & Karim, 2017; Paxton & Glanville, 2015), with some even equating survey trust with general trust (Boyer et al., 2012; Gylfason & Olafsdottir, 2017). While this equation is not rigorous, it reflects researchers' conventional practice of using survey methods to measure general trust. In contrast, researchers use the trust

game to represent both general trust (Dunning et al., 2014; Martini & Torcal, 2019) and specific trust (Anderl et al., 2018; Verberne et al., 2015). Recent research indicates that investment trust is a combination of general and specific trust (Hale et al., 2017). Hale et al. (2017) manipulated both high- and low-trustworthiness virtual characters to play trust games with participants, finding significant positive correlations between investment amounts to both objects, demonstrating general trust characteristics. However, when presenting untrustworthy objects before trustworthy ones, participants' investments differed significantly, showing that trust investment behavior is influenced by trustworthiness and reflects specific trust tendencies. Therefore, the low correlation between survey trust and the trust game may arise because researchers often overlook that the two methods represent different trust types (Holm & Nystedt, 2008; Karlan, 2005). If the measured trust types are matched—both being general trust or the same specific trust—they may show higher positive correlations. Many studies have verified this speculation (Etang et al., 2012; Kovacs et al., 2019; Martini & Torcal, 2019; McEvily et al., 2012). For example, Kovacs et al. (2019) used both survey trust and the trust game to examine Senegalese patients' specific trust in doctors, finding a significant positive correlation between them.

3.3 Limitations of Survey Trust and the Trust Game

From the perspective of paradigm design and application, both survey trust and the trust game have limitations that may contribute to their low correlation. Regarding survey trust, researchers have raised various concerns about scale-based trust measurement (Boyer et al., 2012). First, due to social desirability effects, individuals may misreport their trust levels. For example, research shows that individuals' survey trust levels are often inconsistent with their implicit trust or trust behaviors (Ermisch et al., 2009). Generally, individuals report higher trust levels in surveys than they actually exhibit in behavior (Burns & Conchie, 2008). Second, the concept of the trustee in survey trust often lacks clear definition, and participants may form different interpretations of who constitutes the trustee due to individual and social environmental differences, thereby interfering with results (Abascal & Baldassarri, 2015). For instance, Van Hoorn (2015) showed that influenced by individualistic versus collectivistic cultures, individuals differ significantly in their understanding of the scope of "most people"—that is, their trust radius—when interpreting whether "most people can be trusted." Van Hoorn (2015) found that individualistic cultures have broader trust radii, while collectivistic cultures have relatively narrower ones. Additionally, survey trust suffers from methodological limitations, including potentially imprecise data from subjective ratings, non-standardized data collection methods, lack of economic incentives compared to experimental situations, and insufficient motivation for participants to express their true thoughts (Fehr et al., 2003).

Regarding the trust game, social preferences and risk preferences may influence investment trust levels, causing them to be higher or lower than true trust intentions due to paradigm variations or irrelevant variables, resulting in low cor-

relation with survey trust. Glaeser et al. (2000) found no significant correlation between the trust game and survey trust. Aksoy et al. (2018) speculated that this lack of significant association might be related to the modified paradigm used by Glaeser et al. (2000). Compared to Berg et al.' s (1995) trust game paradigm, Glaeser et al. (2000) eliminated the trustee' s initial endowment and changed investment amounts from continuous to binary. Aksoy et al. (2018) conducted a comparative study of Glaeser et al.' s (2000) modified paradigm and Berg et al.' s (1995) classic paradigm, finding that the classic paradigm showed significant positive correlation with survey trust, while the modified paradigm did not.

In summary, the low correlation between the trust game and survey trust may be related to attitude-behavior differences, measurement type differences, and limitations of both the trust game and survey measures. Therefore, questioning the validity of the trust game based solely on low correlation is clearly debatable.

4 Summary and Outlook

In conclusion, as one of the most widely used trust measurement paradigms, the trust game has its limitations, but current research has not formed a sufficient evidence chain to deny its validity. The main controversies facing the trust game can also be reasonably explained from theoretical perspectives. Meanwhile, numerous studies support the scientific validity of the trust game (Dunning et al., 2014; Corgnet et al., 2016; Javor et al., 2015). Therefore, we argue that under current research conditions, the trust game remains a relatively appropriate method among indirect trust measurement paradigms. The main controversies it faces also serve as driving forces for its continuous development.

4.1 Verifying the Scientific Validity of Trust Game Paradigm Variations

The trust game paradigm has multiple variations, some of which help theoretically avoid the influence of irrelevant variables on research results. For example, changing available investment amounts from continuous to binary helps clarify the influence of social preferences on investment trust levels (Amdur & Schmick, 2012; Ermisch et al., 2009). However, such variations have sparked controversies. For instance, increasing or decreasing available investment amounts may induce different dimensions and levels of risk preferences, thereby affecting trust investment levels (Fairley et al., 2016; Thielmann & Hilbig, 2015). Therefore, extensive empirical research is necessary to verify the scientific validity of trust game paradigm variations. Currently, research in this area is lacking. For example, although Johnson and Mislin (2011) found that available investment amounts significantly affect investment trust levels, their meta-analysis included few studies with large investment amounts, requiring more evidence. Additionally, while the trust game has been widely used in China, systematic research examining the rationality of its paradigm variations remains scarce. Relatedly, the same investment amount may have differential psychological impacts on

participants from different income levels and cultural backgrounds. Therefore, future trust game research should prioritize cross-cultural and cross-regional replication and comparative studies to demonstrate the scientific validity of paradigm variations. Furthermore, considering that social preferences and risk preferences may affect investment trust levels, many researchers have measured them in studies to assess potential errors through statistical means (Corgnet et al., 2016; Breuer et al., 2016; Schlösser et al., 2015). Currently, actively controlling for key irrelevant variables in both experimental design and data processing stages is a beneficial measure for enhancing research validity.

4.2 Further Clarifying the Low Correlation Between Trust Game and Survey Trust

Researchers can further clarify the low correlation issue using the three approaches described above. Additionally, we speculate that the low correlation may also be related to inherent differences in trust types represented by the two paradigms. When using survey methods and trust games without specifying clear trust objects, both default to measuring general trust (Corgnet et al., 2016; Drążkowski et al., 2017; Gereke et al., 2018; Gylfason & Olafsdottir, 2017; Levine & Schweitzer, 2015; Martini & Torcal, 2019; Paxton & Glanville, 2015; Wang & Murnighan, 2017). However, research suggests that when the trust game has no specified investment object, individual investment trust behavior may not represent general trust. For example, Naef and Schupp (2009) simultaneously administered a stranger trust scale (SOEP-trust), a general trust GSS scale, and a trust game without specified investment objects. Results showed that the trust game only significantly correlated with the stranger trust scale, not with the GSS. Similarly, Fehr et al. (2003) found that trust games without specified trust objects significantly correlated with survey trust when the trust object was defined as "strangers." These findings suggest that when trust games do not specify a clear trust object, the trust type represents specific trust toward "strangers" rather than general trust. General trust represents trust in "most people," including both strangers and acquaintances, and is not equivalent to stranger trust. Accordingly, we speculate that when neither survey trust nor trust games specify trust objects, survey trust represents general trust while trust games represent stranger trust, which may contribute to their low correlation.

We further speculate that the closer relationship between unspecified-object trust games and stranger trust may stem from the paradigm' s characteristics. In laboratory settings, whether trust objects are real or virtual, the default assumption when no information about the trustee is provided is that the trustee is a "stranger" to the trustor. Moreover, the trust game simulates real trust scenarios, approximating real-world tasks and including social presses. While evidence is lacking on whether social presses in trust games induce stranger-specific trust, research in other domains shows that when problem-solving measurement paradigms involve real-world tasks, social presses cause individuals

to think more negatively. For example, Kapoor and Khan' s (2019) creativity research found that when creativity problems include real-world contexts, participants exhibit higher levels of negative creativity. Individuals have an automatic vigilance for negative social information (Pratto & John, 1991), and in interpersonal interactions, people tend to identify strangers as negative factors —that is, uncertain and potentially untrustworthy (Feldmanhall et al., 2018). This may make participants more alert to risk factors in interpersonal trust during trust games. Accordingly, trust games may awaken greater attention to negative interpersonal cues, leading participants to interpret trust games with uncertain trustees as stranger games, thereby creating a closer relationship with stranger trust. Of course, this is a bold speculation based on current findings that requires further verification.

4.3 Expanding the Measurement Dimensions of Trust Games

Trust is a multidimensional concept. As previously discussed, inconsistent trust types measured by trust games and survey trust may cause low correlation issues. Compared to survey trust, the trust dimension measured by trust games is very singular, posing certain risks and limitations. For example, trust can be divided into ability trust, integrity trust, and benevolence trust based on trustee characteristics (Mayer et al., 1995). Integrity trust is easily confused with benevolence trust, but they are distinct dimensions with relatively independent mechanisms (Schweitzer et al., 2006). Levine and Schweitzer' s (2015) experiment explored the psychological structure underlying trust games, separating benevolence and integrity components. They found that participants trusted deceptive but altruistic individuals more than honest but selfish ones. This suggests that current trust game paradigms may primarily reflect benevolence-based trust, rarely involving ability and integrity (Levine & Schweitzer, 2015). However, interpersonal trust depends not only on trustees' benevolence but also on their ability and integrity. Therefore, the singular trust dimension representation in trust games not only may cause low correlation with survey trust but also constrains the measurement efficacy of trust games. Future research could attempt to expand trust game measurement dimensions based on Mayer et al.' s (1995) trust theory, starting from how trustee characteristics influence trust dimensions. For example, modifying the trust game scenario so that whether the investment yields benefits depends on the trustee' s knowledge and ability rather than benevolence could represent ability trust.

4.4 Enhancing the Ecological Validity of Trust Games

The trust game simulates real-life interpersonal trust processes and has greater ecological validity advantages compared to survey trust. However, in practice, the game scenarios are somewhat unrealistic. First, in real society, people often face multiple potential trust objects simultaneously, rather than single pairwise games like the prisoner' s dilemma. To address this limitation, some researchers have modified the classic trust game to include one trustor and two trustees, sim-

ulating complex real-world situations with multiple trustees (Zhang et al., 2019). Additionally, traditional trust game paradigms are too abstract for child participants. Some studies have adjusted reward mechanisms for children, such as replacing money with candy and describing complex reward-punishment mechanisms in simple, straightforward interactive game rules, allowing children to experience more authentic trust processes (Chen et al., 2018). Furthermore, traditional trust games are typically conducted in laboratory settings or presented in written form (Gereke et al., 2018; Righetti et al., 2015), with trustees' information often undisclosed (Bauer & Freitag, 2018; Wang & Murnighan, 2017) and no communication between trustors and trustees (Drążkowski et al., 2017). To compensate for these limitations, researchers have used virtual reality technology to further enhance ecological validity. Hale et al. (2017) developed a virtual maze test based on the trust game paradigm. After wearing virtual reality equipment, participants find themselves in a corridor composed of multiple similar rooms. Their task is to choose which rooms to enter based on information presented by virtual characters in the game, ultimately escaping the maze. Researchers comprehensively assess trust levels by recording whether participants consult and follow virtual characters' advice. Virtual reality technology offers good realism, allowing researchers to record more authentic behavioral responses. Future research could develop more ecologically valid trust game tests based on this approach, further enriching trust measurement techniques.

References

Chen, L., Wang, X., Zhao, L., Chen, X., & Wang, Y. (2018). The influence of facial dimorphism on preschool children' s trust behavior: An explanation from personality labels. *Psychological Development and Education*, 34(5), 3-12.

Dou, K., Nie, Y., Wang, Y., & Liu, Y. (2018). Trust or guard? The promoting effect of social mindfulness on cooperative behavior in interactive games. *Psychological Science*, 41(2), 390-396.

Fukuyama, F. (2001). *Trust: The Social Virtues and the Creation of Prosperity.* Hainan Publishing House.

Fu, C., Zhang, Z., He, J., Huang, S., Qiu, J., & Wang, Y. (2018). The dynamic process of general trust game decision-making: Evidence from EEG time-frequency analysis. *Acta Psychologica Sinica*, 50(3), 317-326.

Jiang, S., & Wei, Q. (2013). Beliefs and psychological games: Theory, evidence, and applications. *Economic Research*, 48(6), 141-154.

Liu, G., Xin, Z., & Lin, C. (2017). The bad apple effect and its transmission in interpersonal trust. *Psychology and Behavior Research*, 15(5), 691-696.

Ma, Q., Shen, Q., Li, D., & Wang, K. (2009). Social neuroeconomics: The neurological basis of social decision-making and games. *Journal of Zhejiang University (Humanities and Social Sciences)*, 39(2), 53-63.

Xu, K., & Liu, Y. (2007). Bounded rationality trust view: Integrating rational calculation and irrational attitudes. *Psychological Science*, 30(5), 1137–1143.

Zhang, H., & Wang, E. (2007). Current status and development trends of attitude-behavior relationship research. *Advances in Psychological Science*, 15(01), 163–168.

Zhang, N., Zhang, Y., & Wu, K. (2011). Psychological and neurophysiological mechanisms of trust. *Psychological Science*, 34(5), 1137–1143.

Zhang, X., Huang, D., & Dong, Z. (2019). An experimental study on the impact of information asymmetry on trust behavior in social relationship networks. *Journal of Systems Management*, 28(2), 269–276.

Abascal, M., & Baldassarri, D. (2015). Love thy neighbor? Ethnoracial diversity and trust reexamined. *American Journal of Sociology*, 121(3), 722–782.

Aksoy, B., Harwell, H., Kovaliukaite, A., & Eckel, C. (2018). Measuring trust: A reinvestigation. *Southern Economic Journal*, 84(4), 992–1000.

Amdur, D., & Schmick, E. (2012). Does the direct-response method induce guilt aversion in a trust game? *Economics Bulletin*, 33(1), 687–693.

Anderl, C., Steil, R., Hahn, T., Hitzeroth, P., Reif, A., & Windmann, S. (2018). Reduced reciprocal giving in social anxiety—Evidence from the Trust Game. *Journal of Behavior Therapy and Experimental Psychiatry*, 59, 12–18.

Ashraf, N., Bohnet, I., & Plankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, 9(3), 193–208.

Bauer, P. C., & Freitag, M. (2018). Measuring trust. In E. Uslaner (Eds.), *The Oxford Handbook of Social and Political Trust* (pp. 15–36). Oxford Press.

Bellucci, G., Hahn, T., Deshpande, G., & Krueger, F. (2019). Functional connectivity of specific resting-state networks predicts trust and reciprocity in the trust game. *Cognitive, Affective, & Behavioral Neuroscience*, 19(3), 165–176.

Ben-Ner, A., & Halldorsson, F. (2010). Trusting and trustworthiness: What are they, how to measure them, and what affects them. *Journal of Economic Psychology*, 31, 64–79.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122–142.

Bibi, Z., & Karim, J. (2017). Self-esteem, resilience, and social trust as mediators in the relationship between optimism and job satisfaction: A preliminary analysis of data from European Social Survey. *Pakistan Journal of Psychological Research*, 32(1), 155–173.

Bicchieri, C., Xiao, E., & Muldoon, R. (2011). Trustworthiness is a social norm, but trusting is not. *Politics, Philosophy & Economics*, 10(2), 170–187.

Boyer, P., Lienard, P., & Xu, J. (2012). Cultural differences in investing in others and in the future: Why measuring trust is not enough. *Plos One*, 7(7), 1–5.

Breuer, W., Helduser, C., & Schade, P. (2016). Breaking the rules: Anticipation of norm violation in a binary-choice trust game. *Economics Letters*, 146, 123–125.

Brülhart, M., & Usunier, J. C. (2012). Does the trust game measure trust? *Economics Letters*, 115(1), 20–23.

Bujang, A., Suki, N. M., & Suki, N. M. (2017). Structural relationship between women consumers' attitude and actual behavior in mobile retailing. *Advanced Science Letters*, 23(9), 8946–8949.

Burns, C., & Conchie, S. M. (2008). Measuring implicit trust and automatic attitude activation. In F. Lyon, G. Möllering, & M. N. Saunders (Eds.), *Handbook of Research Methods on Trust* (pp. 239–248). Edward Elgar Publishing.

Carragher, D. J., Thomas, N. A., & Nicholls, M. E. (2018). Is trustworthiness lateralized in the face? Evidence from a trust game. *Laterality: Asymmetries of Body, Brain and Cognition*, 23(1), 1–16.

Castilla, C. (2015). Trust and reciprocity between spouses in India. *American Economic Review*, 105(5), 621–624.

Chetty, R., Hofmeyr, A., Kincaid, H., & Monroe, B. (in press). The trust game does not (only) measure trust: The risk-trust confound revisited. *Journal of Behavioral and Experimental Economics*, 101520.

Cooper, J., & Croyle, R. T. (1984). Attitudes and attitude change. *Annual Review of Psychology*, 35(1), 395–426.

Corgnet, B., Espín, A. M., Hernán-González, R., Kujal, P., & Rassenti, S. (2016). To trust, or not to trust: Cognitive reflection in trust games. *Journal of Behavioral and Experimental Economics*, 64, 20–27.

Derks, J., Lee, N. C., & Krabbendam, L. (2014). Adolescent trust and trustworthiness: Role of gender and social value orientation. *Journal of Adolescence*, 37(8), 1379–1386.

Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution*, 2, 265–279.

Drążkowski, D., Kaczmarek, L. D., & Kashdan, T. B. (2017). Gratitude pays: A weekly gratitude intervention influences monetary decisions, physiological responses, and emotional experiences during a trust-related social interaction. *Personality and Individual Differences*, 110, 148–153.

Dunn, E. W., Aknin, L. B., & Norton, M. I. (2008). Spending money on others promotes happiness. *Science*, 319(5870), 1687–1688.

Dunning, D., Anderson, J. E., 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(1), 122–141.

Ermisch, J., & Gambetta, D. (2006). People' s trust: The design of a survey-based experiment. *IZA Discussion Papers* 2216. Institute for the Study of Labor.

Ermisch, J., Gambetta, D., Laurie, H., Siedler, T., & Noah Uhrig, S. C. (2009). Measuring people' s trust. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(4), 749–769.

Espín, A. M., Exadaktylos, F., & Neyse, L. (2016). Heterogeneous motives in the trust game: A tale of two roles. *Frontiers in Psychology*, 7(1090), 728.

Etang, A., Fielding, D., & Knowles, S. (2012). Are survey measures of trust correlated with experimental trust? Evidence from Cameroon. *The Journal of Development Studies*, 48(12), 1699–1715.

Evans, A. M., & Krueger, J. I. (2011). Elements of trust: Risk and perspective taking. *Journal of Experimental Social Psychology*, 47(1), 171–177.

Evans, A. M., & Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, 42(6), 1585–1593.

Fairley, K., Sanfey, A. G., Vyrastekova, J., & Weitzel, U. (2016). Trust and risk revisited. *Journal of Economic Psychology*, 57, 74–85.

Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association*, 7, 235–266.

Fehr, E., Fischbacher, U., Rosenbladt, B., Schupp, J., & Wagner, G. (2003). A nation-wide laboratory—Examining trust and trustworthiness by integrating experiments in representative surveys. Institute for Empirical Research in Economics, University of Zürich, Working Paper.

FeldmanHall, O., Dunsmoor, J. E., Tompary, A., Hunter, L. E., Todorov, A., & Phelps, E. A. (2018). Stimulus generalization as a mechanism for learning to trust. *Proceedings of the National Academy of Sciences*, 115(7), 1690–1697.

Fetchenhauer, D., & Dunning, D. (2012). Betrayal aversion versus principled trustfulness—How to explain risk avoidance and risky choices in trust games. *Journal of Economic Behavior and Organization*, 81(2), 534–541.

Gächter, S., Herrmann, B., & Thöni, C. (2004). Trust, voluntary cooperation, and socio-economic background: Survey and experimental evidence. *Journal of Economic Behavior & Organization*, 55(4), 505–531.

Gereke, J., Schaub, M., & Baldassarri, D. (2018). Ethnic diversity, poverty, and social trust in Germany: Evidence from a behavioral measure of trust. *Plos One*, 13(7), 1–15.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, 206(2), 169-179.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, 115(3), 811-846.

Gylfason, H. F., & Olafsdottir, K. (2017). Does Gneezy' s cheap talk game measure trust? *Journal of Behavioral and Experimental Economics*, 67, 143-148.

Hale, J., Payne, M. E., Taylor, K. M., Paoletti, D., & Hamilton, A. F. D. C. (2017). The virtual maze: A behavioural tool for measuring trust. *The Quarterly Journal of Experimental Psychology*, 71(4), 989-1008.

Holm, H., & Nystedt, P. (2008). Trust in surveys and games—A methodological contribution on the influence of money and location. *Journal of Economic Psychology*, 29(4), 522-542.

Hong, K., & Bohnet, I. (2007). Status and distrust: The relevance of inequality and betrayal aversion. *Journal of Economic Psychology*, 28, 197-213.

Javor, A., Riedl, R., Kirchmayr, M., Reichenberger, M., & Ransmayr, G. (2015). Trust behavior in Parkinson's disease: Results of a trust game experiment. *BMC Neurology*, 15(1), 126.

Johansson-Stenman, O., Mahmud, M., & Martinsson, P. (2013). Trust, trust games and stated trust: Evidence from rural Bangladesh. *Journal of Economic Behavior & Organization*, 95, 286-298.

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865-889.

Kapoor, H., & Khan, A. (2019). Creativity in context: Presses and task effects in negative creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 13(3), 314-321.

Karlan, D. S. (2005). Using experimental economics to measure social capital and predict financial decisions. *American Economic Review*, 95(5), 1688-1699.

Kerr, M., Stattin, H., & Trost, K. (1999). To know you is to trust you: Parents' trust is rooted in child disclosure of information. *Journal of Adolescence*, 22(6), 737-752.

Kovacs, R. J., Lagarde, M., & Cairns, J. (2019). Measuring patient trust: Comparing measures from a survey and an economic experiment. *Health Economics*, 28(5), 641-652.

Lenton, P., & Mosley, P. (2011). Incentivising trust. *Journal of Economic Psychology*, 32(5), 890-897.

Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes*, 126, 88-106.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.

Martini, S., & Torcal, M. (2019). Trust across political conflicts: Evidence from a survey experiment in divided societies. *Party Politics*, 25(2), 126–139.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734.

McAuliffe, W. H., Forster, D. E., Pedersen, E. J., & McCullough, M. E. (2019). Does cooperation in the laboratory reflect the operation of a broad trait? *European Journal of Personality*, 33(1), 65–83.

McEvily, B., Radzevick, J. R., & Weber, R. A. (2012). Whom do you distrust and how much does it cost? An experiment on the measurement of trust. *Games and Economic Behavior*, 74(1), 285–301.

Michaelidou, N., & Hassan, L. (2014). New advances in attitude and behavioural decision-making models. *Journal of Marketing Management*, 30(5-6), 519–528.

Müller, J., & Schwieren, C. (2019). Big five personality factors in the trust game. *Journal of Business Economics*, 90, 37–55.

Naef, M., & Schupp, J. (2009). Measuring trust: Experiments and surveys in contrast and combination. Working Paper, Royal Holloway College, London.

Ozyilmaz, A., Erdogan, B., & Karaeminogullari, A. (2018). Trust in organization as a moderator of the relationship between self-efficacy and workplace outcomes: A social cognitive theory-based examination. *Journal of Occupational and Organizational Psychology*, 91(1), 221–240.

Paxton, P., & Glanville, J. L. (2015). Is trust rigid or malleable? A laboratory experiment. *Social Psychology Quarterly*, 78(2), 194–204.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32, 122–136.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, 61(3), 380–391.

Reiersen, J. (2019). Drivers of trust and trustworthiness. *International Journal of Social Economics*, 46(1), 2–17.

Righetti, F., Balliet, D., Visserman, M., & Hofmann, W. (2015). Trust and the suppression of emotions during sacrifice in close relationships. *Social Cognition*, 33(5), 505–519.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, 35(4), 651–665.

Rudnicki, K., De Backer, C. J., & Declerck, C. (2019). The effects of celebrity gossip on trust are moderated by prosociality of the gossipers. *Personality and Individual Differences*, 143, 42–46.

Sapienza, P., Toldra-Simats, A., & Zingales, L. (2013). Understanding trust. *The Economic Journal*, 123(573), 1313–1332.

Schlösser, T., Mensching, O., Dunning, D., & Fetchenhauer, D. (2015). Trust and rationality: Shifting normative analyses of risks involving other people versus nature. *Social Cognition*, 33(5), 459–482.

Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes*, 101(1), 1–19.

Shirokova, G., Osiyevskyy, O., & Bogatyreva, K. (2016). Exploring the intention-behavior link in student entrepreneurship: Moderating effects of individual and environmental characteristics. *European Management Journal*, 34(4), 386–399.

Smith, J. R., & Louis, W. R. (2008). Do as we say and as we do: The interplay of descriptive and injunctive group norms in the attitude-behaviour relationship. *British Journal of Social Psychology*, 47(4), 647–666.

Swanson, J. E., Swanson, E., & Greenwald, A. G. (2001). Using the implicit association test to investigate attitude-behaviour consistency for stigmatised behaviour. *Cognition and Emotion*, 15(2), 207–230.

Thielmann, I., & Hilbig, B. E. (2015). Trust: An integrative review from a person-situation perspective. *Review of General Psychology*, 19(3), 249–277.

Uslaner, E. M. (2002). *The Moral Foundations of Trust*. Cambridge University Press.

Van Hoorn, A. (2015). Individualist-collectivist culture and trust radius: A multilevel approach. *Journal of Cross-Cultural Psychology*, 46(2), 269–276.

Verberne, F. M., Ham, J., & Midden, C. J. (2015). Trusting a virtual driver that looks, acts, and thinks like you. *Human Factors*, 57(5), 895–909.

Wang, L., & Murnighan, J. K. (2017). The dynamics of punishment and trust. *Journal of Applied Psychology*, 102(10), 1385–1402.

Woolley, K., & Fishbach, A. (2017). A recipe for friendship: Similar food consumption promotes trust and cooperation. *Journal of Consumer Psychology*, 27(1), 1–10.

Yamagishi, T., Akutsu, S., Cho, K., Inoue, Y., Li, Y., & Matsumoto, Y. (2015). Two-component model of general trust: Predicting behavioral trust from attitudinal trust. *Social Cognition*, 33(5), 436–458.

Note: Figure translations are in progress. See original paper for figures.