

Trust at Zero Acquaintance: More a Matter of Respect Than Expectation of Reward

David Dunning and Joanna E. Anderson
Cornell University

Thomas Schlösser, Daniel Ehlebracht,
and Detlef Fetchenhauer
University of Cologne

Trust is essential for a secure and flourishing social life, but many economic and philosophical approaches argue that rational people should never extend it, in particular to strangers they will never encounter again. Emerging data on the trust game, a laboratory economic exchange, suggests that people trust strangers excessively (i.e., far more than their tolerance for risk and cynical views of their peers should allow). What produces this puzzling “excess” of trust? We argue that people trust due to a norm mandating that they show respect for the other person’s character, presuming the other person has sufficient integrity and goodwill even if they do not believe it privately. Six studies provided converging evidence that decisions to trust follow the logic of norms. Trusting others is what people think they should do, and the emotions associated with fulfilling a social duty or responsibility (e.g., guilt, anxiety) account for at least a significant proportion of the excessive trust observed. Regarding the specific norm in play, trust rates collapse when respect for the other person’s character is eliminated as an issue.

Keywords: trust, trust game, norms, respect, social dilemma

Trust is a paradoxical phenomenon. Consider the fact that trust is crucial for a secure and flourishing social life. It is difficult to think of a marriage or friendship lasting long without trust (Deutsch, 1958; Simpson, 2007) or any organization or firm thriving (Kramer, 1998; Kreps, 1990). It is implausible for a government to survive without trust between its citizens and its political institutions (Fukuyama, 1995; Sullivan & Transue, 1999). Trust is crucial not only among those with established relationships; it is also especially vital between strangers within social groups who have no responsibility toward each other outside of a single, transitory interaction. eBay or the roadside farmer’s market could not exist without trust among strangers. Indeed, trust has been called a core social motive (Fiske, 2003). Nations displaying more trust among strangers tend to have higher rates of economic growth (Fetchenhauer & Van der Vegt, 2001; Knack & Keefer, 1997; Putnam, 1993) and happiness (Oishi, Kesebir, & Diener, 2011).

Albeit essential to social life, trust is also irrational, according to intellectual treatments throughout the centuries. Philosophers such as Machiavelli (1515/2003) and Hobbes (1660/1997) have counseled people against trust. The assertion is that if recipients of trust

are rational actors serving exclusively their own self-interest, they have every reason to exploit trust as soon as it is in their interest to do so. Thus, one should never offer one’s trust unless the other person’s response is placed under such heavy constraints or sanctions that the person is compelled to honor it. In modern times, neoclassical economics most strongly asserts this conclusion (Berg, Dickhaut, & McCabe, 1995; Bolle, 1998). This theoretical inevitability of exploitation makes trust a fool’s errand, especially when dealing with strangers under no compulsion to reward any trust.

Given this contradiction, it is no surprise that trust, especially among strangers, has become a vital topic over the past 20 years in economics (Johnson & Mislin, 2011), sociology (Dunning & Fetchenhauer, 2010a; Snijders & Keren, 2001), political science (Wilson & Eckel, 2011), and even neurobiology (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005). According to the Institute for Science Index, citations to sources mentioning “trust and economic games” or “trust game” (to provide a conservative gauge of interest in trust among strangers) have gone from 6 the year before Berg et al. (1995) published their behavioral measure of trust to 3,947 in 2013.

In contrast, psychology, perhaps the most natural home for trust research, has given the topic comparatively less attention. Although there are extant and important instances of research on intergroup trust (e.g., Brewer, 1999; Dovidio et al., 2008) and interpersonal trust (e.g., Balliet & Van Lange, 2013a; Holmes & Rempel, 1989; Kelley & Thibault, 1978; Mikulincer, 1998), the overall coverage of such a key topic is surprisingly thin, especially given statements that trust is “one of the most important components—and perhaps the most essential ingredient—for the development and maintenance of happy, well-functioning relationships” (Simpson, 2007, p. 587) and “one of the most desired qualities in any close relationship” (Rempel, Holmes, & Zanna, 1985, p. 95).

This article was published Online First May 12, 2014.

David Dunning and Joanna E. Anderson, Department of Psychology, Cornell University; Thomas Schlösser, Daniel Ehlebracht, and Detlef Fetchenhauer, Institute for Sociology and Social Psychology, University of Cologne.

This research was supported financially by National Science Foundation Grant 0745806 awarded to David Dunning and by a Social Sciences and Humanities Research Council postdoctoral fellowship awarded to Joanna E. Anderson.

Correspondence concerning this article should be addressed to David Dunning, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853. E-mail: dad6@cornell.edu

Our goal in the current research is to build on knowledge gleaned from adjacent fields to investigate trust behavior between strangers at zero acquaintance. Despite the protests of philosophers and economists that it is irrational to do so, people trust strangers and those strangers reward that trust in study after study (for reviews, see Balliet & Van Lange, 2013a; Evans & Krueger, 2009; Johnson & Mislin, 2011; Wilson & Eckel, 2011).

We ask what motivates trust in such circumstances. In doing so, we adopt a definition of trust as a psychological state comprising the intention to accept vulnerability based on the chance of reward from positive intentions or behavior of another. This definition is similar to common definitions used in economics and psychology (see Rousseau, Sitkin, Burt, & Camerer, 1998, p. 395), but it contains one major difference: Most theorists view trust behavior as an extension of a person's interpersonal expectations (i.e., people trust others when they expect to be rewarded for it), but readers will soon see we do not hold this view (see Dunning & Fetchenhauer, 2010a, 2010b; Dunning, Fetchenhauer, & Schlösser, 2012; Fetchenhauer & Dunning, 2009).

We begin our analysis by noting that current empirical data present a mystery to explain: People display far too much trust in laboratory interactions between strangers, relative to the rates predicted by current approaches to trust that emphasize risk tolerance and social expectations about whether others will prove trustworthy. We explain this excess of trust by asserting that risk and social expectations fail to play an exclusive role in trust behavior at zero acquaintance. Trust behavior is also importantly influenced by injunctive norms. We hypothesize in particular that people possess a norm to respect another person's character and so trust others to ensure their social behavior aligns with that norm. In their choices, they must give another person the benefit of the doubt when it comes to his or her moral makeup—even if they privately believe that human nature is anything but honorable or benevolent. If they distrust, they risk violating that social pretense.

Assessing Trust via Behavior: The Trust Game

Our assertions about trust at zero acquaintance depend crucially on a disconnection between risk tolerance and social expectations on one side of the ledger and trust behavior on the other. Thus, it is crucial to study trust in a straightforward behavioral context in which risk tolerance, social expectations, and trust behavior can be assessed.

In the current research, we examine trust among strangers with a laboratory paradigm that comes originally from behavioral economics (Berg et al., 1995; Camerer, 2003; Croson & Buchan, 1999; Dunning & Fetchenhauer, 2010b; Kreps, 1990). In this paradigm, alternatively called the *investment* or *trust game*, participants begin with a small amount of money (e.g., \$5), which they are given or requested to bring to the laboratory. Participants are asked whether they wish to keep this money or to give it to an anonymous stranger in the experiment. If they choose to give it, the money is inflated by some factor (e.g., by a factor of 4, to \$20), and recipients of the money are asked whether they wish to keep the entire \$20 or give \$10 back to the first participant. Both players know the rules of the game but not which choice the other player has made. Players are also informed that they will never meet each other and thus will remain forever mutually anonymous.

Variations of this game are common. At times, it has a binary format, as above, in which participants in the role of first person (the trustor) must keep or give all the money in question and those in the role of the second person (the trustee) must keep all or give half back. In other, more continuous variations, either participant can instead give some portion of the money, from 0% to 100%. The amount they give is multiplied by some factor, and the trustee must decide how much, if any, of the resultant amount to give back to the trustor.

At its core, the choice of the first participant mirrors trust as it is theoretically defined: a decision about becoming vulnerable to another person's exploitation to possibly achieve a benefit (e.g., when trusting an acquaintance with a secret entails makes oneself vulnerable to betrayal with the possible benefit of deepening a friendship). Supporting this analogy, the game also fits laypeople's intuitive notions of trust: 78% of respondents describe the game in terms of "trust," "faith in people," or both (Dunning et al., 2012). In other work, when participants judged various labels for the game, the four most-endorsed labels were *Choice*, *Decision*, *Trust*, and *Risk Game* (Anderson & Dunning, n.d.). Further, people subliminally primed with trust concepts give more money to the trustee than do those primed with distrust (Posten, Ockenfels, & Mussweiler, 2013). Additionally, people from countries with high rates of trust in this paradigm tend to endorse trust more on large-scale surveys such as the World Values Survey (Johnson & Mislin, 2012), although the association between survey and behavioral measures of trust is famously weak (Wilson & Eckel, 2011).

Finally and most important, behavior in the game is also correlated with reports about everyday acts of trust, such as loaning money to a friend or leaving a door unlocked (Fehr, Fischbacher, von Rosenblatt, Schupp, & Wagner, 2002; Glaeser, Laibson, Scheinkman, & Soutter, 2000); with measures of generalized trust toward strangers (but not known others; Naef & Schupp, 2009); and with measures of stable individual propensities to trust (Evans & Reville, 2008). Thus, the trust game is well validated as a convenient and controlled way to study trust behaviorally among strangers. In recent years it has come more into the mainstream of social psychology, as researchers have recognized its value as a tool (e.g., Klapwijk & Van Lange, 2009).

Instrumental and Consequentialist Approaches to Trust

Why do people trust despite the cautionary exhortations of economists and philosophers? Ironically, trust makes little sense under a "rational actor" model; yet, virtually all contemporary theoretical treatments of trust in economics and psychology explicitly or implicitly assume the operation of such a rational actor, in that they emphasize an instrumental and consequentialist logic. By instrumental, we mean that people do not trust just for the sake of it but do so in order to gain some future benefit. By consequentialist, we mean that people trust based on the outcomes they think their trust will bring once the trusted person makes a response. That is, trust is simply a means to an end. Under this logic, people trust others when they hold adequate expectations that their trust will be honored and that the payoffs are worth the risk (Berg et al., 1995). Essentially, in the very tradition of ex-

pected utility models, trust is commonly conceptualized as tantamount to expected trustworthiness on the part of others.

Even in psychology, where the notion of people as rational actors has been long since rejected in general, trust is typically studied as the positive outcome of a risk analysis and expectation (Bacharach & Gambetta, 2001; Baldwin, Fehr, Keedian, Seidel, & Thompson, 1993; Barber, 1983; Deutsch, 1958; Evans & Krueger, 2011; Fraley & Shaver, 2000; Kramer & Carnevale, 2001; Murray, Holmes, & Collins, 2006; Pruitt & Rubin, 1986). Indeed, in one recent review of the psychological literature, Simpson (2007) stated that “intergroup researchers, for example, have defined trust as a specific set of socially learned *expectations* [emphasis added] that people hold about various social systems, ranging from other people to social organizations to the larger moral social order” (p. 589). In addition, popular individual difference measures of trust, such as the Rotter scale, focus almost exclusively on generalized expectations about the trustworthiness or benevolence of others (e.g., “Hypocrisy is on the rise in our society”; Rotter, 1967, 1971). General social surveys, such as the World Values Survey, also assess trust by asking about expectations about people in general and a country’s institutions (Johnson & Mislin, 2012).

The Issue: Excessive Trust Given Expectations and Risk Tolerance

Recent experiments using the trust game cause trouble for such accounts of trust. Data repeatedly show that trust fails to follow the instrumental and consequentialist logic assigned to it by theory. People are remarkably cynical about their peers in binary trust games, believing roughly on average that only 45% will hand any money back, even though a large majority—roughly 80%—of those peers typically do (Fetchenhauer & Dunning, 2009). Despite this cynicism, a majority of people still go ahead and trust their peers, with as many as 75% giving money to a stranger despite their pessimistic expectations of return.

They do so even though the risk they take is far greater than the maximum they indicate they are willing to tolerate in other types of gambles with equivalent outcomes. For example, when told explicitly that the chance their trust would be reciprocated was only 46%, the majority (54%) of participants give their money anyway. Significantly fewer (29%) of the same people state a willingness to bet \$5 to possibly win \$10 when presented with the same odds and payoffs in a lottery (Fetchenhauer & Dunning, 2012; see Berg et al., 1995, for similar findings).

Further data show that trust is correlated moderately at best with expectations that the trust will be honored (Fetchenhauer & Dunning, 2009; Glaeser et al., 2000; Lazzarini, Madalozzo, Artes, & Siqueira, 2005) and not at all with a person’s level of risk tolerance (Ashraf, Bohnet, & Plankov, 2006; Ben-Ner & Halldorsson, 2010; Eckel & Wilson, 2004; Kanagaretnam, Mestelman, Nainar, & Shehata, 2009). To be sure, the role of expectation and risk tolerance in trust behavior cannot be totally dismissed, for some researchers find such a relationship (e.g., Evans & Krueger, 2011; Fetchenhauer & Dunning, 2010). In addition, recent reviews of behavior in related social dilemmas, such as the prisoner’s dilemma game, have found strong relationships between social expectations and cooperation (e.g., Balliet & Van Lange, 2013a).

However, when it comes to trust, consequentialist variables such as expectation and risk attitudes at best appear to play only a side

role in the patterns of trust observed. Trust at the behavioral level (e.g., behavioral choice) is not the same as it is at a cognitive level (i.e., beliefs about the trustworthiness of others), and the two should not be conflated. That is, a choice to trust another person should not be automatically taken as a statement of interpersonal optimism at the cognitive level (see Dunning & Fetchenhauer, 2010b). However, if trust at the behavioral level is dissociated from trust at the cognitive level, what explains the dissociation?

An Expressive Account of Trust: On Injunctive Norms

In the following studies, we turn to an account of trust that is expressive in nature rather than instrumental (Dunning & Fetchenhauer, 2010a, 2013; Dunning et al., 2012). By expressive, we mean that trust is motivated by effects produced by the act of trust itself, regardless of more distal outcomes. Many effects derive directly from the act of trust. For example, the action may provide a self-signal of optimism and cooperativeness (Bodner & Prelec, 2002); it may establish a reputation of being a nice and valuable person to others (Akerlof & Kranton, 2000; Fehr, 2004; Lacetera & Macis, 2010); it may also provide the “warm glow” that people feel after prosocial acts (Andreoni, 1989; Dunn, Aknin, & Norton, 2008).

We argue specifically that people trust because the act allows them to honor an injunctive norm. People trust not because it is what they want to do but because they feel it is an obligation of their current social role. It is the action they ought to take. That is, we agree with Deutsch (1973), who said, “to be distrustful is morally far more flagrant than to be credulous. To sin is less virtuous than to be sinned against” (p. 146; see Hartmann, 1932, and Uslaner, 2002, for similar assertions about trust).

Recent work on the trust game demonstrates that behavior is sensitive to normative considerations. Changing labels to components of the game impacts trust rates. For instance, people trust more if they are told they are interacting with a “partner” rather than an “opponent” (Burnham, McCabe, & Smith, 2000). Among the Maasai in Kenya, labeling the game as an *osotua* exercise reduces transfers—presumably because *osotua* signifies a relationship in which individuals or families can make requests freely of each other but only to the extent of their need (Cronk, 2007). Similar labeling effects have been observed in other types of social dilemmas (Batson & Moran, 1999; Kay & Ross, 2003; Liberman, Samuels, & Ross, 2004).

In focusing on normative influence, we should note that our ambition was somewhat different from past research. Rather than looking for moderation of trust by norms, we examined the possibility that norms are largely responsible for producing trust among strangers. That is, if we strip trust behavior down to its basic form, as represented by the standard trust game, does the logic of norms serve to explain the unanticipated “excessive” rate of trust commonly observed? We addressed this question by examining three classic measures of injunctive norms and gauging their relation to trust. First, we asked whether trust follows what people report they should do rather than want to do. Second, we examined whether trust was predicted by emotions associated with honoring social duties and meeting standards of behavior (Higgins, 1987). Finally, we asked how important others would judge one’s behavior in the trust game, with the prediction that respondents

would expect significant others to approve of trust more than distrust (Ajzen & Fishbein, 1980; Fishbein & Ajzen, 2010).

Specifying the Norm

Further, we sought to address two additional questions specifying the type of norm people might follow in their trust behavior.

Social Versus Moral Norms

First, we examined whether the norms associated with trust behavior are social or moral in nature. Careful readers will have noticed that in attributing trust behavior to normative influences, we have steered clear from labeling those normative influences as “social.” We have done so because there is quite a bit of disagreement in the sociological, anthropological, and economic literatures about what constitutes a social norm, relative to more personal or individual ones (Bicchieri, 2006; Dubreuil & Grégoire, 2013; Elster, 2009; Shweder, Much, Mahapatra, & Park, 1997; Turiel, 1983; for a review, see Anderson & Dunning, 2014).

For example, Bicchieri (2006) stated that, for a norm to be social, people must know of its existence and believe that others either follow the norm or think it reflects what one ought to do. Other scholars have gone further, arguing that social norms are vigorously regulated within a cultural community (Elster, 2009): People are rewarded for honoring those norms and punished for violating them, even at cost to the person doing the punishment (Balliet & Van Lange, 2013b; Chudek & Henrich, 2011; Fehr, Fischbacher, & Gächter, 2002; Henrich et al., 2010). Under those definitions, for the norms influencing our participants to be social in nature, we should expect them to operate more strongly when participants’ behavior is public rather than private; that is, when their behavior is open to judgment and sanction.

Norms, however, need not be social, as defined above. Instead, norms can be “moral” in nature, in that they become a standard of behavior that is internalized and autonomous by the time one leaves childhood (see Bicchieri, 2006; Elster, 2009). For instance, norms of equal sharing or helping someone in distress are well internalized by adulthood, thus requiring no further or current social pressure to operate (Elster, 2007). This type of norm is driven simply by an emotional response to a triggering situation, and such responses are not conditioned on whether the behavior is private or public. Thus, our first goal in defining the norm was to identify which of these patterns it follows: that of a social or a moral norm.

Examining the Specific Norm of Respect

We also sought to identify the specific norm people follow when they trust, doing so by obviating it and examining whether trust rates collapsed to levels associated with expectation and risk tolerance. We tested in particular whether people trust out of a norm of respect in which they take care about the signals their trust behavior conveys about the character of the other person. Thus, people may trust to maintain the pretense that the other person has a worthy character; they may avoid distrust because such a choice violates that pretense and shows disrespect. A recurring theme in psychology (Miller, 2006), sociology (Goffman, 1958), and linguistics (Brown & Levinson, 1987) is that people strive to avoid

violating another person’s “definition of a situation,” steering clear of contradicting what they presume the other person believes about themselves and the world around them. In particular, people respect the “positive face” of the other person; namely, the pretense that the other person has the character of a rational, moral, honorable person who acts in good faith and with goodwill (Brown & Levinson, 1987).

Research on speech acts reflects this respect: People respond to sexist views with deflection and distraction rather than confrontation (Swim & Hyers, 1999; Woodzicka & LaFrance, 2001), accept apologies they privately believe to be insincere (Risen & Gilovich, 2007), and avoid displaying disagreement with conversation partners (DePaulo & Bell, 1996). Herein, we suggest this desire to maintain social pretenses extends from speech to the realm of behavior. To trust is to affirm positive face. To distrust instead is a “face threat” in that it questions the moral character of the other individual. Note that when people apply this logic, their actual beliefs about whether the other person is honorable are irrelevant. Their actions, regardless of belief, must maintain that pretense.

Overview of Studies

The six studies that follow examined three general questions. Studies 1 through 4 examined whether trust behavior displays the signature of a norm and also whether that norm is more social or moral in nature. Studies 5 and 6 examined whether the specific norm people follow is one of respect for the other person’s character: Would trust rates deflate if the issue of the other person’s character were removed from the decision?

We also examined whether trust behavior would once again fail to follow the logic of expectation and risk tolerance (Fetchnauer & Dunning, 2009, 2012). We predicted that too many people would trust, given their (pessimistic) expectations about others trustworthiness and their level of risk tolerance. We predicted that expectations of reciprocation and risk tolerance would at best be weak and unreliable predictors of who trusts versus who does not.

Study 1: Probing for Indications of an Injunctive Norm

In Study 1, we examined whether decisions to trust followed the logic of injunctive norms. We presented participants with a version of the binary trust game described above, in which the initial endowment is \$5, multiplied to \$20 if given to Person B, and explicitly asked participants what they “should” do as well as what they “wanted” to do. In doing so, we were able to investigate whether impressions of social obligation or correctness explain why so many people trust relative to the rate suggested by an economic analysis. We predicted that, regardless of what people wanted to do, participants overall would feel they should trust rather than not trust, and that this tendency would at least partially explain why people trust more than the rate suggested by risk tolerance and expectations.

We also examined which emotions—those related to social obligation and duty versus those related to desire—best predicted why “too many” people trust. Much work, in particular research on self-discrepancy theory (Higgins, 1987), shows that concerns about social obligation bring about feelings of agitation—namely, anxiety and guilt—when people contemplate failing to live up to

social standards (Higgins, Bond, Klein, & Strauman, 1986; Higgins, Shah, & Friedman, 1997; Roney & Sorrentino, 1995). In contrast, emotions associated with desire tend to focus on what would bring contentment and pleasure into a person's life (Haselhuhn & Mellers, 2005; Mellers, 2000; Mellers, Schwartz, & Ritov, 1999).

A few extant studies provide some supportive, albeit indirect, evidence that trust behavior involves more agitation emotions than desire-related ones. First, when people report the emotions they feel as they contemplate trust, guilt and calm are better predictors of trust than are other emotions (Schlösser, Fetschenhauer, & Dunning, 2013). In addition, when people are induced to feel fear (vs. anger), they bet less on a lottery but risk more on another person in a trust game (Kugler, Connolly, & Ordóñez, 2012). This is consistent with agitation emotions nudging people toward adhering to a social obligation or norm in the trust game.

As such, we proposed that participants would report more agitation (i.e., anxiety and guilt) when considering forgoing trust than giving it, and that this overall bias across participants would, in part, explain why people trust to a degree beyond the limit imposed by their expectations and risk tolerance.

Method

Participants. Participants were 38 Cornell University students from large lecture psychology courses who took part for extra credit toward their course grades.

Procedure. Participants were approached at the end of class lectures and asked if they wanted to stay to take part in an experiment for extra credit. Those who did were told that they were to make a series of decisions (labeled merely "Decision 1," "Decision 2," etc.) about what to do with money and that one of the decisions would be a real one, to be carried out with actual money at the end of the study. They were informed that they would not learn which decision they was the real one until they had made all decisions, and they were admonished to take all decisions seriously. Finally, participants were also instructed that their decisions would be anonymous, such that even the experimenter would not know which decisions they reached.

In all, participants considered three scenarios. The first, a measure of risk tolerance, asked them to consider a chance to gamble \$5 to win \$10 from the draw of a ball from an urn. Participants were told that the urn contained 100 balls and were asked to state the minimum number of winning balls the urn would have to contain to entice them to gamble the \$5.

In the second scenario, participants were asked to consider the binary form of the trust game. They were told that they had been matched to another individual in the experiment but that this person would remain unknown to them throughout (and after) the experiment. They were told they could be assigned the role of either Person A or B with this individual. If they were assigned to be Person A, they would have to decide whether they wanted to keep \$5 (in which case Person B would receive nothing) or give the \$5 to Person B. If they gave the money to Person B, the experimenter would inflate the amount to \$20 and give Person B a choice: to either give \$10 back to Person A or keep the entire \$20. Person B would have to make this decision before knowing what Person A had decided.

Participants were first asked to consider what they *wanted* to do and *should* do in this situation, with the additional instruction that "sometimes, what people should do and what they want to do lead to the same choice, sometimes they lead to different choices." Participants were then presented with two bipolar 7-point scales, one asking them to rate how much they wanted to keep the money versus give the money and the other asking them to report how much they should keep the money versus give it. Higher numbers indicated that one wanted to or should give rather than keep. The midpoint of the scale (4) was labeled as "I am indifferent or ambivalent." Order of the two scales was counterbalanced across participants.

Participants also rated how choosing to give versus keep the \$5 would make them feel. They rated their feelings related to giving along 14 different emotion dimensions and did the same for their feelings about keeping the money. Three emotions had been classified a priori as related to wanting and desire (*content*, *happy*, and *pleased*; α s = .70 and .81 for keeping and giving, respectively). Four were classified as agitation emotions relating to violating social obligations and standards (*anxious*, *guilt*, *remorse*, and *tense*; α s = .89 and .72 for keeping and giving, respectively). Next, participants indicated which decision they would make as Person A and also indicated which decision they would make as Person B. In addition, they estimated the overall percentage of participants in the study who would choose to give money back versus keep the entire \$20.

The final scenario that participants faced was a simple one: whether to receive \$5 with certainty or to bet that \$5 to possibly win \$10 on a flip of a coin. After participants completed this decision and placed their questionnaires in anonymous folders, they were told that this final decision was the one that would be carried out. For those who decided to gamble, the experimenter flipped a coin to determine each participant's payment separately (with "heads" the winning side of the coin). Other participants were given \$5.

Results and Discussion

Of 38 participants, 27 (71%) decided to hand their \$5 to Person B in the trust game.

Consequentialist predictors. This high level of trust emerged although participants, on average, thought that only 52.5% ($SD = 22.8\%$) of their peers would return money but required, on average, a 63.8% ($SD = 14.6\%$) chance of winning in order to induce them to gamble \$5 on the first question about risk tolerance. In fact, according to their reports of risk tolerance and expectations of their peers, only 14 participants (37%) should rationally have decided to trust, in that their expectations that their peers would return the \$10 matched or exceeded their risk tolerance to gamble \$5 on the previous question about the urn. Indeed, 16 (42%) participants gave money to Person B despite a risk tolerance and peer expectation that would predict keeping it, whereas only 3 (8%) participants displayed the opposite pattern. Thus, replicating past work, far too many participants trusted relative to the predictions made by typical, expectation-based definitions of trust ($p = .004$, two-tailed, by sign test). Further demonstrating that trust decisions were not driven by expectations or simple risk aversion, estimates of the percentage of peers who would return money and risk tolerance

both failed to predict trust behavior (see Table 1). In short and echoing past research, trust decisions did not appear to emerge as a function of a simple expectation-risk tolerance calculation.

We also found no evidence that a descriptive norm played any role in the excess of trust rates. As a measure of descriptive norm on average, participants thought that only 48.6% ($SD = 18.90$) of their peers would trust. Thus, there was no overall perception that an overwhelming majority of peers would trust, potentially putting pressure on participants themselves to trust to excess. What, then, did predict why so many people trusted?

Should versus want. Not surprisingly, wanting to trust and feeling one should trust were both related to the decision participants made. Ratings of “want” and “should” were uncorrelated with each other, $r(36) = .14, p = .40$, but each in separate logistic regressions predicted the likelihood that a participant would trust (see Table 1). Both variables were significantly related to trust when placed in a simultaneous regression (for want, $L-R \chi^2 = 21.00, p < .001, OR_{\text{standardized}} = 21.00$; for should, $L-R \chi^2 = 7.89, p = .005, OR_{\text{standardized}} = 7.95$).¹ Should ratings remained significant after we controlled for expectations that Person B would reciprocate trust (see Table 1).

However, further analyses revealed that only should ratings could explain why so many participants trusted over and above the economic analysis. Overall, ratings of want ($M = 3.97, SD = 2.11$) fell almost exactly on the scale midpoint of indifference (i.e., 4) between trusting and not trusting ($t < 1, ns$), with virtually equal numbers stating they wanted to trust versus not trust (17 vs. 18, respectively). Thus, there was no overall bias in what people wanted to do that could explain why the majority opted for trust over the alternative.

However, for ratings of should, there was a clear bias. Participants on average thought they should trust the other person over forgoing trust ($M = 5.50, SD = 1.81$), which was significantly different from the scale midpoint of indifference (see Table 1). Of our 38 participants, 29 (76%) stated they should trust, whereas only 6 (16%) stated the opposite. Thus, there was a sizable consensus among our participants that they should trust. According to our logistic regression, if participants overall had been indifferent between giving versus keeping the \$4 in their should ratings (i.e., the mean rating fell at the scale midpoint of 4), the trust rate would have dropped from 71%, the observed rate, to 55%.

Agitation versus contentment. An examination of the emotions participants attached to keeping versus giving the money also showed that both desire and norms anticipated their decisions. We took the difference in ratings of the desire emotions that participants reported when they considered trusting versus not trusting. We did the same for the social obligation emotions related to agitation. In separate regression analyses, both differences predicted the likelihood of trust. The more contented participants thought they would be with a decision to trust over forgoing trust, the more they opted to give the \$5; the more agitated they felt about trusting versus not, the less participants opted to trust (see Table 1). The relation of agitation to trust decisions remained significant even after controlling for participants' expectations about Person B (see Table 1).

Again, it was agitation and not contentment that was the most likely candidate to explain why an excess of participants trusted. Participants were more content to trust (vs. not trust) but to a degree only barely reaching marginal significance ($M_s = 3.12$ and

2.67, $SD_s = 0.94$ and 0.80 , respectively; see Table 1). In contrast, participants were much more agitated about forgoing trust than trusting ($M_s = 2.69$ and $1.85, SD_s = 1.41$ and 1.04 , respectively; see Table 1). In terms of the logistic regression, if participants had been, instead, equally agitated by both options, trust rates would have fallen to 64%. (A comparable analysis for contentment emotions lowers trust rates from 71% down to only 69%.)²

Summary. In sum, in Study 1 we found that trust followed the logic of injunctive norms much more than consequentialist, economic principles. Consequentialist variables, such as risk tolerance or expectations of reward, failed to predict trust decisions. Indeed, far too many people trusted given their relative aversion to risk and cynicism they held about the trustworthiness of their peers. Although over 70% of participants trusted, the rate predicted purely by consequentialist variables was a mere 37%.

Only the pressure of a norm could explain why so many participants chose to trust. First, participants, overall, felt that they should trust their peer, and second, they reported more guilt and anxiety when they contemplated not doing so. In short, two indicators of feelings of obligation best predicted trust behavior, thus suggesting the guiding hand of a norm in people's choices.

Study 2: Further Indications

In Study 2 we further tested for normative influences on trust behavior. We assessed subjective injunctive norms using a measure adapted from the planned theory of behavior (Ajzen & Fishbein, 1980; Fishbein & Ajzen, 2010), which probes for injunctive norms by asking about the perceived approval of significant others for specific behaviors. Targets whose approval we measured included formative influences on participants' lives; namely, one's mother and father, as well as currently influential peers, such as one's friends and fellow students. Parents are instruments of primary socialization, modeling for their children appropriate norms, attitudes, and behaviors (Bronfenbrenner, 1970; Higgins, 1989; Maccoby & Martin, 1983); thus, reports of their approval may reflect a particularly ingrained injunctive norm. The approval of friends and fellow students—participants' current peer group and secondary instruments of socialization—presumably represents an injunctive norm for how peers feel one ought to behave.

¹ In all logistic analyses, we report likelihood-ratio χ^2 statistics rather than Wald statistics because the former is the recommended test to ascertain the contribution of any variable toward predicting an outcome (Hosmer & Lemeshow, 2000).

² Another way to depict these emotion results is to conduct a 2 (Decision: trust vs. not) \times Option (Emotion attached to: trusting vs. not) mixed model analyses of variance (ANOVAs), with the last variable serving as a within-subject variable. An analysis of agitation emotions showed that participants who trusted were more agitated about keeping the money ($M = 3.00$) than they were about giving it ($M = 1.82$), $t(36) = 4.05, p < .001, d = 0.78$. Participants who kept the \$5 felt equally agitated about the two options ($M_s = 1.93$ and 1.91 for keeping and giving, respectively), $t(38) = 0.05, p = .961, d = 0.02$. This pattern gave rise to an overall main effect for decision option, in which participants felt more agitated about keeping the \$5 than giving it, $F(1, 36) = 4.93, p = .033, \eta^2 = .11$, as well as a Decision \times Option interaction, $F(1, 36) = 4.56, p = .039, \eta^2 = .10$. For contentment emotions, a significant Decision \times Option interaction arose, $F(1, 36) = 6.28, p = .017, \eta^2 = .15$ ($M_s = 3.18$ and 2.67 for keeping and giving, respectively, for those who kept; $M_s = 2.46$ and 3.31 , respectively, for those who gave), but there was no overall bias toward wanting to give the \$5, $F(1, 36) = 0.38, p = .541, \eta^2 = .01$.

Table 1
Biases Toward or Against Trust Among Various Predictors Along With Their Correlation to Choosing Trust (Study 1)

| Measure | Directional bias toward trust | | | | Correlation with trust | | |
|---|-------------------------------|------------------|----------|----------|------------------------|--------------|----------|
| | Raw difference | Cohen's <i>d</i> | <i>t</i> | <i>p</i> | Standardized <i>OR</i> | L-R χ^2 | <i>p</i> |
| Consequentialist predictors | | | | | | | |
| Expectation of reciprocation | | | | | 1.49 | 1.17 | .278 |
| Risk tolerance | | | | | 1.20 | 0.26 | .610 |
| Descriptive norm ^a | -1.60 | -0.08 | -0.45 | <.001 | 37.15 | 23.62 | <.001 |
| Injunctive norm and preference predictors | | | | | | | |
| Should trust ^b | 1.50 | 0.83 | 5.10 | <.001 | 2.65 | 7.00 | .008 |
| Want to trust ^b | -0.03 | -0.01 | -0.08 | .939 | 16.30 | 21.11 | <.001 |
| Agitation ^c | -0.84 | -0.53 | -3.88 | <.001 | 0.42 | 4.63 | .032 |
| Contentment ^c | 0.46 | 0.28 | 1.72 | .094 | 3.46 | 6.91 | .009 |
| Controlling for expectation | | | | | | | |
| Should trust | | | | | 2.54 | 5.83 | <.001 |
| Agitation | | | | | 0.44 | -5.14 | <.001 |

Note. L-R χ^2 = likelihood-ratio chi-square; *OR* = odds ratio.

^a Relative to 50%. ^b Relative to scale midpoint. ^c Ratings for trust minus forgoing trust.

Note that examining perceptions of parents and peers separately allowed us to begin to tease apart whether the norms operating were social or moral in nature. The logic of social norms suggests that perceptions of both peers and parents would matter, in that a social norm is one that is believed to be actively regulated within one's current social environment. The logic of a moral norm would suggest that parents would matter but not peers. That is, it is one's socialization in early years that matters. Active regulation in one's current social environment would not be relevant (Bicchieri, 2006; Dubreuil & Grégoire, 2013; Elster, 2009).

We also sought to replicate the emotion findings from Study 1, expecting agitation emotions to better predict an excess of trust relative to desire emotions.

Method

Participants. Participants were 72 Cornell University students who participated for extra credit toward their psychology course grades. They participated in small groups or singly.

Procedure. As in Study 1, participants were told that the study focused on financial decision making and that they were to make a series of decisions about what to do with money, one of which would be carried out for real. They were also instructed that their decisions would be anonymous, and even the experimenter would not know which decisions they made.

The series of decisions that participants confronted was identical to that in Study 1. After making these decisions, participants completed a 10-item measure of norms adapted from Ajzen and Fishbein (1980; Fishbein & Ajzen, 2010). They rated the approval of five different targets separately for the actions of keeping and giving the \$5 in the trust game; targets included the participants' mothers, fathers, friends, and most Cornell students (e.g., "Most of my friends would approve if I chose Option 1 [keeping the \$5]"). Ratings were made on 7-point scales anchored by *definitely would disapprove* to *definitely would approve*. For the purpose of analysis, we took the mean of the parents' approval ratings for each option ($r_s = .64$ and $.42$ between mother and father approval, $p_s < .001$, for keeping and giving, respectively). We also took the mean of the peers'—friends' and fellow students'—approval ratings for

each option ($r_s = .34$ and $.31$, $p_s < .01$, for keeping and giving, respectively).

Participants also rated how choosing to give versus keep the \$5 would make them feel, in measures similar to those used in Study 1. They separately rated their feelings related to giving versus keeping the money along 12 emotion dimensions. Three were again classified a priori as related to wanting and desire (*content*, *happy*, and *pleased*; $\alpha_s = .87$ and $.86$ for keeping and giving, respectively). Four were classified as agitation emotions related to violating social obligations and standards (*anxious*, *guilty*, *remorseful*, and *tense*; $\alpha_s = .80$ and $.55$ for keeping and giving, respectively).

Finally, participants made several other economic decisions, including—as in Study 1—whether to receive \$5 with certainty or to bet that \$5 on the chance of winning \$10 on a flip of a coin. This last choice was the decision that they played for real.

Results and Discussion

Of the 72 participants in the study, 1 who did not complete the crucial trust decision was excluded, leaving 71 for analysis. Thirty-seven (52%) chose to give their \$5 in the trust game.

Consequentialist predictors of trust. Replicating Study 1, trust behavior failed to follow consequentialist logic. On average, participants thought that 40.1% ($SD = 20.8\%$) of their peers would return money while requiring, on average, a 66.4% ($SD = 13.1\%$) chance of winning to gamble \$5 to possibly win \$10 on the risk tolerance question about balls in an urn. Thus, even more striking than in Study 1, only 10 participants (14%) should have decided to trust. However, 30 (42%) gave money to Person B even though such a decision contradicted what their risk tolerance and peer expectation reports would have suggested; only 3 (4%) participants displayed the opposite pattern. Thus, as in Study 1, far too many participants trusted relative to the strict predictions made by an economic analysis ($p < .001$, by sign test).

To examine consequentialist predictors of behavior on an individual level, we conducted separate logistic regressions to predict the odds of participants deciding to trust from participants' risk tolerance and expectations for their peers' behavior. Risk tolerance

was not a significant predictor of behavior. However, expectations for peers' behavior significantly predicted trust behavior (see Table 2).

Social approval. On average, participants thought their parents (see Table 2) would be significantly more approving of giving \$5 ($M = 3.72, SD = 0.76$) than keeping it ($M = 3.21, SD = 0.95$). In contrast, there was no significant difference between how approving participants thought their peers would be of giving \$5 ($M = 3.50, SD = 0.79$) versus keeping it ($M = 3.32, SD = 0.83$). Thus, participants appear to hold an injunctive norm—a norm stating that they should trust another person, as indicated by perceptions of their parents, but not one that is shared their peers.

Separate logistic regression analyses (see Table 2) revealed that parental approval predicted the odds that someone would trust. According to this analysis, if participants had thought that their parents showed no bias toward trust, their trust rate would have dropped from 52% to 43%. A further analysis showed that parent approval continued to predict trust even after controlling for expectation of trustworthiness (see Table 2). Perceptions of peer approval, in contrast, failed to predict trust rates (L-R $\chi^2 = 0.36, p = .55, OR_{\text{standardized}} = 1.15$).

Finally, we found no evidence for a descriptive norm supporting an excess of decisions to trust. Although participants' decisions to trust were correlated strongly with the percentage of peers they believed would trust, they also thought that only 37.0% of their peers would choose to trust, which suggests, if anything, an anti-trusting descriptive norm.

Agitation versus contentment. At the mean level, participants were significantly more likely to say they would feel agitated (see Table 2) if they kept the \$5 ($M = 2.66, SD = 0.96$) than if they gave it to Person B ($M = 2.19, SD = 0.80$). In contrast, they exhibited no consensus about which option would make them feel more contentment-related emotions ($M_s = 2.83$ and $3.05, SD_s = 0.87$ and 0.91 , for keeping and giving, respectively).

To examine the impact of emotions, we fit logistic regression models to predict the odds of deciding to trust. Agitation emotions significantly predicted trust rates (see Table 2), and the analysis suggested that the trust rate would have dropped to 42% if participants had felt no more agitated about keeping the money than they did about giving it. The link between agitation and trust was significant even after controlling for expectations about Person B. A separate logistic regression found that contentment-related emotions also predicted trust rates, also doing so after expectations were controlled (L-R $\chi^2 = 10.58, p = .001, OR_{\text{standardized}} = 2.51$).³

Summary. In sum, we found that trust was again predicted by signs of injunctive norms much more strongly than by a rational economic logic. Although participants' social expectations about reciprocity of trust predicted their decisions, too many participants trusted, given those expectations. This surplus in trust was accounted for by people's emotions—specifically those relating to agitation at the thought of not trusting—and the perceived approval of influential models of normative behavior (i.e., their parents). Approval from peers did not matter, giving initial evidence that the norms surrounding trust are more moral than social in nature.

Study 3: Expectations Attached to Individuals

The first two studies suggested that trust decisions were driven more by an injunctive norm than by expectations of reward. However, participants reported their expectations about their peers in general, and that may not be the expectation that matters in driving trust. Instead, the crucial expectations may be those attached to specific other individuals. According to this logic, if we had given participants a reason to expect trustworthiness or its opposite from the specific person they were paired with, those expectancies would have driven trust decisions and swamped concerns about norms. Indeed, people make vigorous inferences about the trustworthiness of others from the instant they meet, from simple facial features (Olivola & Todorov, 2010; Oosterhof & Todorov, 2009; Willis & Todorov, 2006). Why make such strong inferences if not to guide behavior?

In Study 3, we tested this account by going one step beyond zero acquaintance to give participants information about the individual they were paired with a photo that displayed that other person's face. We then examined whether the expectations of trustworthiness inspired by those faces would swamp the role played by norm-related variables.

Method

Participants. Participants were 72 Cornell University students who participated for extra credit toward their psychology course grades. They participated in small groups or singly.

Procedure. Participants received the same instructions as in Study 1. The sole procedural addition was providing a photograph of the purported Person B on the page of the questionnaire describing the trust game. Participants were randomly assigned to receive 1 of 8 faces, of which 4 had been rated as trustworthy and 4 less so. In a pilot study, a group of 53 respondents had rated 23 college-age faces along a 5-point scale. The 4 faces rated most trustworthy ($M_s = 3.42$ – $3.60, SD_s = 0.79$ – 0.88) and the 4 rated least trustworthy ($M_s = 2.30$ – $2.62, SD_s = 0.69$ – 0.94) differed significantly from each other in all pairwise comparisons, $t_s(52) > 5.09, p_s < .001$.

In addition to completing measures completed in Study 1, participants estimated the likelihood that their specific Person B would give money back. They also completed a 10-item scale assessing their overall liking of Person B. After completing these counterbalanced measures, participants indicated their decisions as Person A and Person B. They also completed two other economic

³ As in Study 1, a 2 (Decision: trust vs. not) \times Option (Emotion attached to: trusting vs. not) mixed model ANOVA on agitation emotions showed that participants who trusted were more agitated about keeping the money ($M = 2.99$) than they were about giving it ($M = 2.04$), $t(36) = 5.43, p < .001, d = 0.89$. Participants who kept the \$5 felt equally agitated about the two options ($M_s = 2.30$ and 2.35 for keeping and giving, respectively), $t(33) = -0.28, p = .781, d = -0.05$, giving rise to an overall main effect for decision option where participants felt more agitated about keeping the \$5 than giving it, $F(1, 69) = 12.64, p < .001, \eta^2 = .13$, as well as a Decision \times Option interaction, $F(1, 36) = 15.69, p < .001, \eta^2 = .16$. For contentment emotions, a significant Decision \times Option interaction arose, $F(1, 36) = 13.02, p < .001, \eta^2 = .15$ ($M_s = 2.89$ and 2.59 for keeping and giving, respectively, for those who kept; $M_s = 2.77$ and 3.47 , respectively, for those who gave), but there was no overall bias toward wanting to give the \$5, $F(1, 36) = 2.04, p = .162, \eta^2 = .02$.

Table 2
Biases Toward or Against Trust Among Various Predictors Along With Their Correlation to Choosing Trust (Study 2)

| Measure | Directional bias toward trust | | | | Correlation with trust | | |
|-------------------------------|-------------------------------|------------------|----------|----------|------------------------|--------------|----------|
| | Raw difference | Cohen's <i>d</i> | <i>t</i> | <i>p</i> | Standardized <i>OR</i> | L-R χ^2 | <i>p</i> |
| Consequentialist predictors | | | | | | | |
| Expectation of reciprocation | | | | | 1.68 | 4.24 | .039 |
| Risk tolerance | | | | | 1.21 | 0.63 | .430 |
| Descriptive norm ^a | -13.00 | -0.66 | -5.54 | <.001 | 3.32 | 17.08 | <.001 |
| Social approval | | | | | | | |
| Parental ^b | 0.51 | 0.38 | 3.16 | .002 | 3.11 | 14.77 | .008 |
| Peer ^b | 0.18 | 0.13 | 1.07 | .287 | 1.15 | 0.36 | .564 |
| Person B ^b | 3.42 | 4.56 | 38.48 | <.001 | 1.20 | 0.57 | .450 |
| Emotion | | | | | | | |
| Agitation ^b | -0.84 | -0.40 | -3.39 | .002 | 0.33 | 14.81 | .032 |
| Contentment ^b | 0.22 | 0.17 | 1.46 | .148 | 2.59 | 12.25 | .009 |
| Controlling for expectation | | | | | | | |
| Parental approval | | | | | 3.02 | 5.83 | <.001 |
| Agitation | | | | | 0.34 | -5.14 | <.001 |

Note. L-R χ^2 = likelihood-ratio chi-square; *OR* = odds ratio.

^a Relative to 50%. ^b Ratings for trust minus forgoing trust.

decisions: the lottery and the coin flip, as in the first two studies. As usual, the coin flip decision was the one that they played for real at the end of the study.

Results and Discussion

The overall trust rate in the study was 65.3%. Liking measures did not predict trust decisions nor the results reported below.

Consequentialist predictors. As intended, expectations about Person B were influenced by the photo manipulation. Participants viewing a high-trustworthy photo estimated the likelihood of receiving money back to be significantly higher than those viewing a low-trustworthy photo, $t(70) = 3.31, p = .002, d = 0.76$ ($M_s = 55.6\%, SD = 22.9\%$, vs. $39.2\%, SD = 19.1\%$, respectively). However, the photo manipulation failed to affect estimates of the overall percentage of participants who would share the \$20 with Person B ($M = 53.2\%, SD = 24.8\%$ vs. $M = 46.6\%, SD = 21.0\%$ vs. in the high and low conditions, respectively), $t(70) = 1.22, p = .23, d = 0.29$.

Despite the success of the manipulation, trust behavior again failed to follow consequentialist logic. Based on their reports of risk tolerance and expectations that Person B would return money, 45% of participants in the high-trust condition should have given the money to Person B, whereas only 17% of participants in the low-trust condition should have done so: a significant difference, $\chi^2(1, 72) = 6.74, p = .017$, odds ratio (*OR*) = 4.00. However, in reality, trust rates in the two conditions were virtually identical, 64.5% and 65.9% in the high- versus low-trustworthy conditions, respectively, $\chi^2(1, 72) = 0.01, p = 1.00, OR = .94$.

Across the two photo conditions, we again saw an excess of trust. In the high-trustworthy condition, 29% of participants gave money to Person B even though such a decision contradicted what their risk tolerance and peer expectation reports would have suggested. In the low-trustworthy condition, that figure was 54%, with only 10% and 5% of participants per condition, respectively, displayed the opposite pattern. Thus, again, far too many participants overall trusted relative to the predictions derived from a purely consequentialist economic analysis ($p < .001$, by sign test).

To examine consequentialist predictors of behavior on an individual level, we conducted a logistic regression analysis to predict the odds of participants deciding to trust from specific expectations of Person B (see Table 3). The analysis revealed a significant relationship but one weaker than that of trust decisions with general estimates of trustworthiness. Risk tolerance, again, failed to correlate with trust rates.

We also found no evidence for the influence of a descriptive norm on decisions to trust. To be sure, participants who estimated a higher proportion of their peers would choose to trust were themselves more likely to trust (L-R $\chi^2 = 14.87, p < .001, OR_{\text{standardized}} = 3.21$). But, on average, participants thought that significantly less than half of their peers would choose to trust ($M = 41.0\%, SD = 21.2\%$), $t(71) = -3.59, p = .001, d = 0.42$, so descriptive norms cannot account for the high mean level of trust observed.

Should versus want. Ratings of “want” and “should” were moderately correlated in this study, $r(72) = .33, p = .004$, but neither differed by condition ($F_s < 1.4, p_s < .25$). Should and want ratings independently predicted the likelihood that a participant would trust (see Table 3). Both these relationships emerged unchanged when entered simultaneously in a logistic regression analysis along with photo condition and all possible interactions. The relationship between should and trust remained significant in an additional models in which we controlled either for expectations based on Person B's face (L-R $\chi^2 = 13.92, p < .001, OR_{\text{standardized}} = 3.20$) or people in general (L-R $\chi^2 = 12.41, p < .001, OR_{\text{standardized}} = 2.72$).

Should ratings explained again why so many participants trusted over and above the economic analysis. Overall, ratings of want ($M = 3.58, SD = 1.99$) fell marginally below the scale midpoint of indifference (i.e., 4); that is, on the side of not trusting (see Table 3). Thus, if anything, this slight bias in what people wanted would predict the majority opting to keep the money rather than trust Person B. However, for ratings of should, there was a clear bias in favor of trusting ($M = 5.26, SD = 1.88$), significantly higher than the scale midpoint (see Table 3). The logistic regres-

Table 3
Biases Toward or Against Trust Among Various Predictors Along With Their Correlation to Choosing Trust (Study 3)

| Measure | Directional bias toward trust | | | | Correlation with trust | | |
|---|-------------------------------|------------------|----------|----------|------------------------|--------------|----------|
| | Raw difference | Cohen's <i>d</i> | <i>t</i> | <i>p</i> | Standardized <i>OR</i> | L-R χ^2 | <i>p</i> |
| Consequentialist predictors | | | | | | | |
| Expectations of reciprocity based on | | | | | | | |
| Specific face | | | | | 1.90 | 5.89 | .015 |
| People in general | | | | | 2.47 | 11.24 | <.001 |
| Risk tolerance | | | | | 1.27 | 0.94 | .330 |
| Descriptive norm ^a | -9.00 | -0.42 | -3.59 | <.001 | 3.20 | 14.87 | <.001 |
| Injunction norm and preference predictors | | | | | | | |
| Should trust ^b | 1.26 | 0.67 | 5.69 | <.001 | 2.59 | 12.45 | <.001 |
| Want to trust ^b | -0.42 | -0.21 | -1.78 | .080 | 7.47 | 33.50 | <.001 |
| Agitation ^c | -0.73 | -0.58 | -4.90 | <.001 | 0.40 | 10.38 | .001 |
| Contentment ^c | 0.11 | 0.12 | 1.03 | .300 | 1.05 | 0.04 | .841 |

Note. L-R χ^2 = likelihood-ratio chi-square; *OR* = odds ratio.

^a Relative to 50%. ^b Relative to scale midpoint. ^c Ratings for trust minus forgoing trust.

sion analysis predicted that the trust rate would have dropped from 65% to 52% if should ratings had showed no bias toward the decision to trust.

Agitation. As in Study 1, participants were significantly more likely to say they would feel agitated if they kept the \$5 ($M = 2.85$, $SD = 1.00$) than if they gave it to Person B ($M = 2.12$, $SD = 0.97$). Photo condition had no impact on emotion ratings. Agitation about giving the money relative to keeping it up significantly predicted trust rates (see Table 3) and was not influenced by photo condition in a separate analysis. Similar to the case of should ratings, without this bias in agitation emotions, the predicted rate of trust according to a logistic regression analysis would have deflated to around 55%. The link between agitation and trust was significant even after controlling for expectations about Person B (L-R $\chi^2 = 10.83$, $p = .001$, $OR_{\text{standardized}} = 0.34$) or people in general (L-R $\chi^2 = 10.64$, $p = .001$, $OR_{\text{standardized}} = 0.38$).⁴

Summary. In sum, we found that having more information about a stranger with whom one is interacting changes some features of the situation (e.g., by influencing one's belief that the stranger will repay one's trust and one's liking of that person). However, the central findings of our previous studies replicated in spite of these changes. This pattern of finding lends further support to our hypothesis that trust behavior with strangers is driven by normative concerns.

Study 4: Trust in Public Versus Private

The primary goal in this study was to examine the social versus moral norm distinction. According to most current definitions, for a norm to be social, people must, first, be aware of it and, second, conform their behavior to it more when in a public setting where they are subject to social pressure. For instance, they may feel a need to maintain their reputation among their peers as generous and considerate individuals and so trust more when their behavior is public (Fehr, 2004; Henrich, 2006; Milinski, Semmann, & Krambeck, 2002). To test this account, we varied whether participants made their decisions in complete privacy versus announced them to three other participants and the experimenter. If people trust to enhance their reputation, they should trust more when their decisions are made public rather than kept private. If, instead, they

trust to conform to a moral norm, the public versus private nature of their decision should not matter (Bicchieri, 2006; Dubreuil & Grégoire, 2013; Elster, 2009).

In addition, often in behavioral economics, the trust game is not as a binary choice but rather a continuous decision in which people can give some portion of their money, up to 100%, to the other person. Any portion they decide to give is multiplied by some factor, and the recipient is given the option of returning any amount of the money he or she receives (from zero to all of it) to the participant (Berg et al., 1995; Croson & Buchan, 1999). Study 4 presented participants with a continuous version of the trust game, in which participants were given €5 and asked if they wished to give any or all of it to an anonymous recipient. Any amount given was multiplied by 4, and the recipient was then given the choice to return any amount they saw fit. Thus, our second goal on Study 4 was to replicate the findings of Study 1 with respect to feeling that one should trust other people: As participants were considering their decisions, they reported the amount they wanted to and should give, and we examined whether the amount they transferred to the recipient better matched what they wanted versus felt they should transfer.

Method

Participants. Participants assigned to the role of Person A were 136 students from the University of Cologne. An additional 198 students were recruited separately for the role of Person B at a later time.

⁴ Again, a mixed model ANOVA on agitation emotions showed that trusting participants were more agitated about keeping the money ($M = 3.04$) than they were about giving it ($M = 1.97$), $t(45) = 5.91$, $p < .001$, $d = 0.87$. Participants keeping the \$5 felt equally agitated about the two options ($M_s = 2.50$ and 2.39 for keeping and giving, respectively), $t(24) = 0.51$, $p = .617$, $d = 0.10$, giving rise to an overall main effect for decision option and Decision Made \times Option Considered interaction, $F(1, 69) = 10.75$, $p = .002$, $\eta^2 = .11$. In addition, overall, participants were more agitated about keeping than giving the \$5, $F(1, 69) = 16.21$, $p < .001$, $\eta^2 = .16$. For contentment emotions, a Decision Made \times Option Considered mixed model ANOVA produced no significant effects, in contrast to previous studies.

Procedure. Upon arriving at the laboratory, Person A participants received €5 and were asked whether they would give some portion of it to another student (Person B) who would remain anonymous. They could give any amount from €0 to all €5 in €1 increments. That amount was multiplied by 4, and Person B would then decide whether to return any portion of the money. Participants reported the amount of money they personally wanted to give, the amount they should give, and then the amount they would actually give.

We varied whether participants made their decisions privately or publicly. In the private condition, participants reported their decision on a questionnaire placed in an envelope identified only by a password known to the participant. As such, neither the experimenter nor Person B knew the identity of any participant. In the public condition, participants read their decision out loud to three other participants in the same role in the experiment, as well as the experimenter, just after making their decision.

Person A participants were then matched with 136 recipients from the separately recruited second pool of Person B participants. Each of the latter was told to consider all possible contributions from Person A (e.g., €1, €2, all the way to €5) and indicated, in €1 units, how much they wished to return to Person A after that contribution had been multiplied by 4. Recipients were then given the amount they had retained for themselves. The experimenters returned to the initial set of Person A participants 2 weeks later to give them the amount that their Person B had designated via an envelope containing that amount.

Results and Discussion

Once again, trust decisions were a mixture of desire and injunctive norm. The amount participants gave to the other person correlated with both what they wanted to do, $r(134) = .38$, $p < .001$, and what they felt they should do, $r(134) = .25$, $p = .003$. Want and should ratings were negatively correlated with each other, $r(134) = -.21$, $p = .014$.

These correlations, however, underplay the strength of injunctive norms (i.e., should) in predicting trust decisions. Overall, the amount of money participants gave more closely resembled their reports of what they should do than what they wanted to do. Participants thought they should give more money ($M = €3.54$, $SD = 1.98$) than they wanted to ($M = €2.45$, $SD = 2.10$), $t(135) = 4.02$, $p < .001$, $d = 0.35$. The actual amount they gave ($M = €3.69$, $SD = 1.67$) was far above what they reported wanting to give, $t(135) = 6.80$, $p < .001$, $d = 0.59$, but not significantly different from what they felt they should give ($t = 0.76$, $p = .45$, $d = 0.07$).

Considered another way, only 33% of participants wanted to give all their money, but 58% felt they should. The latter percentage failed to exceed the proportion (53%) of those who ultimately did give all their money (*ns* by sign test) to the recipient but was far more than the proportion of those who professed to want to ($p < .001$ by sign test). By contrast, 33% of participants wanted to give nothing, but only 19% felt they should give nothing to the recipient. The actual percentage giving nothing was 10%, a rate lower than that saying they should ($p = .029$ by sign test) or wanted ($p < .001$) to give nothing.

Did making decisions public have any impact on trust rates? Making them public made participants want to give more to the

recipient, relative to those in the private condition ($M_s = €2.91$ and €2.00, respectively, $SD_{pooled} = 2.06$), $t(134) = 2.57$, $p = .011$, $d = 0.44$. However, making the decision public failed to influence what people thought they should give ($M_s = €3.48$ and €3.61, for public and private conditions, respectively, $SD_{pooled} = 1.99$, $t = -0.38$, $p = .70$, $d = 0.07$), and the amount they ultimately gave ($M_s = €3.85$ and €3.54, for public and private conditions, respectively, $SD_{pooled} = 1.66$), $t(134) = 1.10$, $p = .27$, $d = 0.19$. Publicity also failed to influence the percentage of participants who gave all €5 (57% vs. 49% in public and private conditions, respectively) or none (11% vs. 9% in public and private conditions, respectively), both chi-squares < 1 , $ps > .39$, $ORs < 1.35$.

Incidentally, Person's B on average gave back €0.36, €1.94, €3.87, €5.83, and €7.93 for contributions from Person A of €1, €2, €3, €4, and €5, respectively. Thus, Person A nearly "broke even" if giving €2 and made a profit for larger amounts. Replicating past findings (Pillutla, Malhotra, & Murnighan, 2003), Person B was more generous as the amount contributed by Person A rose. They returned only 9% of the total money they had available when Person A contributed €1 but nearly 40% when Person A gave all €5.

In sum, Study 4 provided further evidence in support of the importance of norms for predicting trust decisions. In a continuous version of the trust game, participants gave amounts that better matched the level they thought they should send as opposed to the level they stated they wanted to send. In addition, this study furnished evidence against a reputational explanation of overtrusting: Participants were no more likely to trust when their decision was public than when it was private, despite our using a more sensitive continuous measure of trust (rather than a cruder binary one). Although any null effect must be interpreted with caution, this finding suggests that people's trust decisions were not driven by a desire to look good to others. Together with the social approval data from Study 2, these data suggest that the norm involved in trust decisions is more moral and personal in nature than social.

Study 5: Respecting the Other's Character

In Study 5, we turned to identifying the specific injunctive norm people were honoring with their trust decisions. There are many to consider. Are people trying to avoid seeming selfish (Monin & Miller, 2001)? Are they trying to create potentially equal (i.e., fair) payoffs to themselves and the other person (Bolton & Ockenfels, 2000)? Are they trying to increase the total amount of money available to the world (Hardin, 1968)?

We assert that people trust to avoid signaling their doubts about the other person's moral character. They likely presume that others see themselves as reasonable, honorable, and capable individuals. Thus, via their actions, people must choose whether to respect what they presume those others think of themselves by trusting them or to insult those others by withholding trust (Brown & Levinson, 1987; Goffman, 1958).

We tested this assertion by comparing behavior in the trust game with an alternative game that precludes any signal about the other person's perceived trustworthiness. Specifically, Person A was told that if he or she handed over the \$5, Person B would then flip a coin to determine whether or not to return \$10. Thus, the outcome was not a product of Person B's character but of chance

instead. We predicted that people would feel less of a social obligation or norm to give the \$5 with Person B's character no longer at stake, and thus they would do so significantly less often than they would in the standard trust game.

Method

Participants. Participants were 88 Cornell University undergraduates who took part for extra credit toward their course grades. Of those, 51 were recruited from large lecture psychology courses; the rest participated in small groups or singly in the lab.

Procedure. Participants first completed the risk tolerance measure used in the first three studies and then were randomly assigned to consider one of two games: the binary trust game used in the previous studies, or a variation called the *coin flip game*. The coin flip game differed from the trust game only in that Person B could not choose how to divide the money. Instead, Person B would have to flip a coin to determine whether he or she would keep the entire \$20 or split it evenly with Person A. Participants made the same estimates and should and want ratings as in Study 1 for whichever game they were assigned to.

Participants then estimated the percentage of Persons A who would choose to give the \$5 to Person B; only those in the trust game condition estimated the percentage of Person Bs who would choose to give money back versus keep the entire \$20. They then indicated which decision they would make as Person A and which decision they would make as Person B. Finally, participants faced the decision of whether or not to bet \$5 on the flip of a coin for a chance at \$10; this was, as usual, the decision that they played for real.

To confirm our intuition that giving Person B a choice was more polite and respectful of the character of Person B, we collected pilot data from 21 additional respondents, recruited from Amazon.com's Mechanical Turk crowdsourcing facility (who were paid \$0.30 for their responses). Respondents in this pilot had the three options for Person A described to them: keeping the \$5; giving the \$5 to Person B and letting that person make a choice whether to give money back; or giving the money to Person B, who then would flip a coin to see if any money would be given back. Respondents then rated the characteristics of each option on 5-point scales, with end points labeled 1 (*not at all*) and 5 (*very much so*). As seen in Table 4, respondents rated giving Person B a choice as the option with the most social decorum. It was rated

as significantly more polite and less insulting than directing Person B to flip a coin or keeping the money. This choice was also seen as showing more respect than they other two options. In terms of commenting on Person B's character, respondents rated trust as showing less doubt and more confidence in Person B, and less likely to signal a suspicion that Person B was greedy, relative to the other two options of requesting a coin flip or keeping the money.

Results and Discussion

Of the 88 participants in the study, 2 who had participated in other trust game studies in the past were excluded, leaving 86 for these analyses.

Impact of game on choice. A greater percentage of participants gave their money to Person B in the trust game condition (30 of 45, or 66.7%) than in the coin flip condition (18 of 41, or 43.9%), $\chi^2(1) = 4.51, p = .033, OR = 2.55$. This pattern emerged although participants in the trust game condition thought, on average, that there was only a 36.7% ($SD = 29.1$) chance of receiving money back, a figure significantly less than the 50% chance of return found in the coin flip condition, $t(44) = -3.07, p = .004, d = 0.46$.

The role of injunctive norms. Further analyses suggested that behavior differed between the trust and coin flip games because the former evoked a norm that the latter did not. Participants confronting the trust game stated that they should give the \$5 to a greater degree than did those in the coin flip game ($M_s = 5.36$ and 4.39, respectively, $SD_{pooled} = 2.06$), $t(84) = 2.17, p = .033, d = 0.47$. Should ratings in the trust game condition were also significantly biased toward giving the \$5 compared to scales midpoint, $t(44) = 4.44, p < .001, d = 0.66$, but not in the coin flip variant, $t(40) = 1.21, ns, d = 0.19$. Across the two conditions, should ratings predicted rates of giving the \$5 (L-R $\chi^2 = 17.27, p < .001, OR_{standardized} = 2.66$). In terms of "want," participants across the two conditions did not significantly differ ($M_s = 3.58$ and 4.00 for trust and coin flip conditions, respectively, $SD_{pooled} = 2.15$), $t(84) = -0.91, p = .37, d = 0.20$; nor did participants show a bias in either condition toward wanting to give versus keep the money ($t_s > -1.23$).

Several tests indicated that should ratings mediated the link between type of game and willingness to gamble the \$5. As seen in Figure 1, type of game predicted how much participants thought

Table 4
Ratings of Choices in Three-Option Trust Game Along Dimensions of Decorum and Character

| Measure | Choice | | | <i>t</i> of difference between trust and flip | <i>t</i> of difference between trust and keep |
|--|--------|-----------|----------|---|---|
| | Trust | Flip coin | Keep \$5 | | |
| Is respectful of Person B | 4.10 | 3.10 | 1.52 | 3.09 | 6.53 |
| Is polite | 4.19 | 3.19 | 1.52 | 3.87 | 6.24 |
| Is rude or insulting | 4.73 | 3.81 | 1.95 | 4.16 | 6.42 |
| Overall decorum | 4.33 | 3.37 | 1.67 | 5.27 | 6.83 |
| Suggests some doubts about the character of Person B | 4.14 | 2.86 | 1.95 | 3.45 | 4.65 |
| Suspects Person B is selfish or greedy | 4.38 | 3.29 | 1.71 | 3.32 | 8.00 |
| Places confidence in the attributes of Person B | 4.38 | 2.33 | 1.48 | 5.77 | 6.21 |
| Overall character | 4.30 | 2.83 | 1.71 | 5.55 | 7.50 |

Note. All $ps < .001$, degrees of freedom (dfs) = 20.

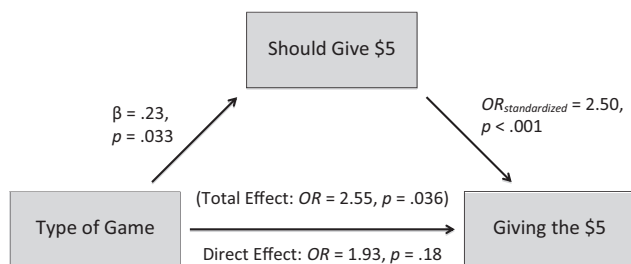


Figure 1. Mediation analysis of trust game versus coin flip game choices in Study 5. *OR* = odds ratio.

they should give the \$5, and should ratings predicted giving the \$5 even after statistically controlling for type of game. Type of game significantly influenced rates at which participants gave the \$5 initially but not after controlling for should ratings. All these results point to injunctive norms, as represented by should ratings, successfully mediating the impact of game type (trust vs. coin flip) on the rate at which participants gave the \$5 to Person B.

In addition, to more directly assess whether an injunctive norm mediated the impact of the game on rates of handing the money over, we conducted a bootstrap analysis with 5,000 iterations with condition as the predictor (trust game = 0, coin flip = 1), decision as the outcome (keep = 0, give = 1), and should ratings as the mediator (using the Preacher & Hayes, 2008, method for dichotomous outcomes). This analysis revealed a standardized indirect effect, 95% CI [-1.073, -0.056]. Because the confidence interval did not cross 0, this analysis reveals a significant indirect effect of game on ultimate decisions that flowed through what participants thought they should do. When we tested want ratings and estimates of the percentage of Person As who would give the \$5 as mediators instead, confidence intervals for both indirect effects crossed 0, indicating no mediational role.

Summary. This study revealed that the decision to give money to another person was more likely when that decision involved trusting the other person's judgment rather than the vagaries of a coin flip, despite the better perceived odds of reward in the coin flip condition. This finding suggests that signaling regard for another person's character is an important normative component to the decision to trust: People felt more strongly that they should give the money when any reward depended on the judgment of the other person rather than a coin flip, and those perceptions of what they should do explained the higher rating of giving the \$5 seen in the trust game versus the coin flip variant.

Study 6: Respect Redux

In Study 6 we sought to replicate the role played by respect by pitting the choices from Study 6 against each other directly in a within-subjects format. That is, participants in this study were presented a trust game with three options, the usual two (trust or keep) plus an option to give the \$5 to Person B and compel that person to flip a coin to determine whether the money was split. Again, if the norm directing behavior in the previous studies operates to maintain the pretense of another person's good character, participants should opt for the decision that gives Person B a choice over the one that gives no choice (flip a coin), irrespective of expectations of that other person's reciprocity.

As in previous studies, we asked participants what they should do versus wanted to do. We also asked them to rate each decision option along attributes related to being intelligent and rational as well as the norm of honoring another person's character (i.e., politeness, respect).

Method

Participants. Forty-one Cornell University students participated for extra credit toward a course grade.

Procedure. Participants followed the procedure of Study 1, except for the following changes. First, they were given three options in the trust game: keep the \$5, trust the \$5 to Person B to choose between keeping and splitting it, or give the \$5 to Person B to compel a flip a coin to determine whether Person B kept it or split it with Person A. The last two options were counterbalanced in order. In addition, before providing their decision in the trust game, participants rated each of the three options separately in terms of what they wanted to do versus should do using 7-point Likert scales. They also rated the qualities of each option on several dimensions using 7-point scales: rational/irrational, intelligent/stupid, polite/insulting, respectful/disrespectful, greedy/generous, and unselfish/greedy. The first two ratings were combined as an index of rationality; the second two as an index of respect, and the final two as an index of altruism.

Results and Discussion

Participants showed a clear preference among the three options. The majority, 22 participants (54%), opted to give Person B the \$5 and give that person a choice about what to do next. A total of 10 participants (24%) opted to keep the money; only 9 (22%) opted to give the \$5 and compel Person B to flip the coin, $\chi^2(2) = 7.66, p = .006$. Thus, even in a direct comparison of options, the norm to trust another person's judgment beat a coin flip alternative, $\chi^2(1) = 5.40, p = .020, d = 0.46$, which was equally altruistic in terms of increasing wealth for self and other. Participants chose trust more often even though on average they thought that only 42.8% ($SD = 25.41$) of peers would return money if given a choice, a figure marginally lower than the 50-50 odds in the coin flip, $t(40) = -1.80, p = .079, d = 0.28$. Expectation of having trust reciprocated did predict choosing trust over the other two options (L-R $\chi^2 = 4.54, p = .033, OR_{\text{standardized}} = 2.06$). However, even those choosing trust thought, on average, that only 50.4% of their peers would honor it (versus an average of 20.7% for those keeping the money and 38.9% for those opting for the coin flip).

Injunctive norms. Further analyses again suggested that it was an injunctive norm that produced the excess of trust. First, as Table 5 shows, participants felt they should trust more than they thought they should keep the \$5, $t(40) = 2.65, p = .012, d = .42$, or opt for the coin flip, $t(40) = 3.19, p = .003, d = .50$. Next, we constructed an index of should ratings for trust versus the other two options (i.e., $2 \times [\text{should rating for trust}] - [\text{should rating for keep}] - [\text{should rating for coin flip}]$) and used this difference score to predict trust behavior over the other options in a logistic regression (L-R $\chi^2 = 5.92, p = .015, OR_{\text{standardized}} = 2.32$). According to this analysis, if people were indifferent between trust and the other two choices on this measure of injunctive norm, only 43%

Table 5
Overall Ratings of Person A's Options in Study 6

| Attribute | Option | | | <i>F</i> | <i>p</i> |
|---------------|---------------------|-------------------|-------------------|----------|----------|
| | Trust | Keep money | Flip coin | | |
| Should trust | 5.12 _a | 3.73 _b | 3.76 _b | 5.72 | .007 |
| Want to trust | 4.27 _{a,b} | 4.54 _a | 3.41 _b | 3.57 | .038 |
| Respectful | 5.70 _a | 3.08 _b | 4.91 _c | 32.81 | <.001 |
| Rational | 3.53 _a | 4.85 _b | 3.90 _a | 4.47 | .018 |
| Altruistic | 5.83 _a | 2.36 _b | 5.06 _c | 44.55 | <.001 |

Note. *d*'s are 2 and 39 for should and want ratings and are 2 and 38 for respectful, rational, and altruistic ratings. Means in a row not sharing the same subscript are significantly different from each other ($p < .05$) by simple paired *t* test not corrected for multiple comparisons.

would have opted for trust. In a supplemental analysis controlling for expectations, the link between should ratings and trust remained significant (L-R $\chi^2 = 5.46$, $p = .020$, $OR_{\text{standardized}} = 2.40$).

Ratings of want did not explain the excess of trust. To be sure, an index of wanting to trust over the other two options (constructed using the same method described above) predicted who would trust over one of the other options (L-R $\chi^2 = 22.94$, $p < .001$, $OR_{\text{standardized}} = 9.56$). However, as seen in Table 5, people on average wanted to keep the money just as much as they wanted to trust, $t(40) = 0.49$, $p = .63$, $d = .08$. They also wanted to trust only marginally more than they wanted to flip the coin, $t(40) = 1.94$, $p = .060$, $d = .30$. Thus, overall, participants did not want to trust more than they wanted to choose one of the other alternatives. That is, the index of wanting was not significantly positive overall ($M = 0.59$, $SD = 5.63$), $t(40) = .66$, $p = .51$, $d = .10$. This failure to find a bias in wanting toward trust meant that wanting was not a plausible candidate to explain why so many people opted to trust over the other options available.

Respect and other option attributes. Recall that we also asked participants to rate each choice along three other dimensions: respect, rationality, and altruism. We predicted that perceptions of the trust option as respectful and polite would explain the excess of trust decisions. Indeed, Table 5 shows that trust was seen as more respectful than keeping the money, $t(39) = 8.07$, $p < .001$, $d = 1.28$, or choosing the coin flip, $t(39) = 3.55$, $p < .001$, $d = 0.57$. An index of respect (i.e., $2 \times$ [respect rating for trust] – [respect rating for keep] – [respect rating for coin flip]) predicted who chose trust over the alternatives in a logistic regression at a rate just missing statistical significance (L-R $\chi^2 = 3.77$, $p = .052$, $OR_{\text{standardized}} = 2.04$). This analysis suggested that if participants had thought trust to be only as respectful as the other two alternatives, only 35.5% would have chosen it. When controlling for expectations, the link between respect and trust became statistically significant (L-R $\chi^2 = 4.48$, $p = .034$, $OR_{\text{standardized}} = 2.62$).

Ratings along the other attributes did not explain the excess in trust. In terms of rationality, participants thought that keeping the money was the smart thing to do. Participants rated this option more favorably along this attribute than either of the other two options (t 's > 2.5 , p 's $< .02$, d 's $> .40$; see Table 5). Participants rated trust as the most altruistic choice, even over flipping the coin (t 's > 3.4 , p 's $< .005$, d 's > 1.35), but an altruism index, constructed like those above, failed to predict who would trust versus choose

another alternative (L-R $\chi^2 = 2.41$, $p = .25$, $OR_{\text{standardized}} = 1.72$). Finally, there was no descriptive norm supporting trust rates. On average, participants thought that 30.4% of their peers would choose trust, a figure not different from the proportion thought to keep the money (38.5%) or choose the coin flip (31.0%), $F(2, 39) = 1.60$, $p = .21$, $\eta^2 = .08$.

Summary. In sum, Study 6 implicated injunctive norms as an explanation for the excess of trust observed in all five studies. Trusting another person proved to be the most popular option over keeping the money or directing the other person to respond by flipping a coin. Again, trust was seen as the act that people should do, and this perception predicted who would trust versus not. More specifically, the action of trust was seen as more respectful and less insulting than the two alternatives—a perception that also predicted who would trust.

Participants also viewed trust as the most altruistic act of the three, although this impression did not predict who would trust. It appears that participants based their perceptions of which act was selfless and generous on more than simply the amount of wealth an option could generate, for trust and the coin flip produced equal increases in the amount of money in play. Rather, participants also based these judgments on procedural grounds; that is, on the method used to distribute that wealth. Granting choice to the other person was not only respectful, it was considered selfless. This finding harkens back to classic justice research on the distinction between distributive (i.e., who gets the money) and procedural justice (i.e., the procedure by which the distribution is decided; Lind, 1988), and it serves as a reminder that people may not only be interested in material features of a situation but also its interpersonal meaning.

General Discussion

Trust is essential for civilized society and thriving interpersonal relationships, yet these studies demonstrate that trust fails to follow the logic typically assigned to it by theory. Theorists from both economics and psychology assume that people will trust when they are reasonably optimistic that others will prove trustworthy; that is, at a level within their tolerance for risk. However, study after study has shown that trust behavior among strangers fails to align with risk attitudes or comport with people's expectations of reward (Fetchnauer & Dunning, 2009, 2012). The result is too much trust at zero acquaintance, given an analysis based on consequentialist logic. The six studies reported herein replicated this effect. Across the four studies that presented a simple binary trust game (Studies 1–3, Study 5, trust game condition), a full 62.4% of participants (141 out of 226) chose to trust when their expressed levels of risk tolerance and expectations about the trustworthiness of their peers suggested that only 20.4% (46) would accept the gamble.

A Meta-Analytic Review of Possible Mechanisms

A meta-analysis of all the data from those four studies converges on an explanation of this excessive trust: namely, trust is neither completely instrumental nor consequentialist in nature. Instead, trust is at least in part an expressive act. People are concerned about the act itself and what it represents or immediately brings about.

More specifically, people trust because they are concerned about an injunctive norm. They trusted so often because it was what they felt they should do, not necessarily what they wanted to do. Across the four studies featuring a binary trust game and a measure of whether people felt they should trust (including the parental approval measure of Study 2), participants showed a clear bias that they should trust Person B rather than forgo trust, and such responses were significantly correlated with the decision to go ahead and trust (see Table 6). People showed no comparable bias toward wanting to trust Person B. In addition, participants across three of those studies indicated that trusting, relative to not trusting, would leave them feeling less agitated (e.g., nervous, tense)—an emotional indicator that people were concerned about obeying norms—and such feelings were again correlated with decisions to trust. The correlations of what one should do and agitation emotions were still significant even after controlling for expectations that the other person would reciprocate trust.

In addition, a meta-analysis provided some evidence for and against other alternative accounts of trust. For example, our studies showed that descriptive norms did not explain why so many people trusted. In fact, people tended to trust even though they thought it a rare behavior (see Table 6). To be sure, participants who trusted tended to think that more of their peers would trust, but they also thought the descriptive norm among their peers was not to trust.

However, there was sufficient evidence in the combined data to suggest that trust might arise, at least in part, due to the warm emotional glow of prosocial behavior (Andreoni, 1989; Dunn et al., 2008). Specifically, participants across the four studies associated the decision to trust with contentment-related emotions (e.g., pleased, content) more than they did the alternative of forgoing trust, and those feelings did predict who trusted. However, in terms of statistical strength, the link between trust and “warm glow” emotions did not reach the levels associated with social obligation emotions such as guilt and anxiety (see Table 6).

Finally, our data do not totally dismiss a role for social expectation. In the meta-analysis, expectations that the other person would reciprocate trust predicted who would trust (see Table 6). However, although expectations moderated trust rates, they in no

way produced them, in that a majority of participants trusted even though only a small minority held expectations and risk attitudes that should have allowed for that trust.

The Normative Influence: Social or Moral?

Beyond the meta-analysis, our data specified that participants showed many signs of following a more internal and private norm in their trust behavior than they did one that was more social, as defined by several recent conceptual treatments of norms across several behavioral and social sciences (Bicchieri, 2006; Dubreuil & Grégoire, 2013; Elster, 2009; Shweder et al., 1997; Turiel, 1983).

In particular, the behavior of our participants failed to follow the signature of a norm that was social. According to many theorists, for a norm to be social, people must explicitly believe such a norm exists and that it is actively enforced. Our participants did not believe in a norm to trust. Instead, they believed the opposite: that a majority of their peers would forgo trust rather than extend it. In addition, participants in Study 2 did not believe their peers harbored an overall opinion that one should trust, much less an opinion they would enforce. Finally, in Study 3, making trust decisions open to public inspection did not inspire more trust. These results are consistent with the data of other researchers, who find that individuals fail to believe others will punish those who withhold trust, only those who fail to reciprocate it (Bicchieri, Xiao, & Muldoon, 2011).

Thus, the norm participants followed in the trust game appears to be, as Bicchieri (2006) termed it, moral in nature (see also Elster, 2009). Trusting another person follows a moral standard that is internal to the individual and that requires no social pressure to evoke. Consistent with this analysis, the act of trust is associated with emotional and cognitive responses connected to fulfilling one’s social duty (see Studies 1 and 2), and stripping an action of its injunctive nature (i.e., the coin flip decisions of Studies 4 and 5) is sufficient to dampen people’s tendency to take a risk on another person.

Table 6

Meta-Analysis of Biases Toward or Against Trust Among Various Predictors Along With Their Correlation to Choosing Trust

| Measure | Directional bias toward trust | | | Correlation with trust | | |
|--|-------------------------------|---------------------|----------|------------------------|---------------------|----------|
| | <i>d</i> | Stouffer’s <i>Z</i> | <i>p</i> | Standardized <i>OR</i> | Stouffer’s <i>Z</i> | <i>p</i> |
| Injunctive norm and preference predictors | | | | | | |
| Should trust ^a | 0.63 | 8.23 | <.001 | 2.62 | 5.57 | <.001 |
| Want to trust ^b | −0.13 | −1.74 | .081 | 9.55 | 8.22 | <.001 |
| Agitation ^c | −0.50 | −6.30 | <.001 | 0.38 | −4.84 | <.001 |
| Contentment ^c | 0.19 | 2.39 | .017 | 2.37 | 3.75 | <.001 |
| Controlling for expectation of reciprocation | | | | | | |
| Should trust | | | | 2.70 | 5.83 | <.001 |
| Agitation | | | | 0.39 | −5.14 | <.001 |
| Consequentialist predictors | | | | | | |
| Expectation of reciprocation ^d | | | | 2.21 | 5.02 | <.001 |
| Risk tolerance | | | | 1.11 | 0.72 | .41 |
| Descriptive norm | −0.40 | 5.84 | <.001 | 12.72 | 8.49 | <.001 |

Note. Includes data from Studies 1, 2, 3, and 5 (trust game condition only). *OR* = odds ratio.

^a Including parental approval measure of Study 2. ^b Not measured in Study 2. ^c Not measured in Study 5. ^d Includes general population expectations from Study 3.

Taken together, these results suggest that there is a rich vein of research questions to explore about social versus moral norms in cooperative and prosocial behavior. Currently, there is vigorous theorizing that people help and cooperate each other due to active enforcement of publicly known social norms (Balliet & Van Lange, 2013b; Chudek & Henrich, 2011; Fehr et al., 2002; Henrich et al., 2010). Our results complicate this theorizing by showing that public enforcement and social sanction need not always be in place for people to act in a trusting way. More internal norms are in play, and the question becomes which acts are more internal versus social and how do those that are more internal (i.e., moral) develop?

Respect

In addition, the last two studies suggested the specific norm that participants were following. Participants wished to avoid showing disrespect for the moral character of their interaction partner. Thus, when we removed the issue of the partner's character by making the partner's response depend on a coin flip rather than on an intentional choice, participants were much less willing to make themselves vulnerable to the other person. Study 5 showed this in a between-subjects format; Study 6 did so in a within-subjects format. In Study 5, perceptions of what one should do, rather than what one wanted to do, mediated the different rates of gambling on another person found in the trust versus the coin flip game. In Study 6, ratings of what should do and what was the most respectful choice to make correlated with greater rates of choosing to trust another person's decision rather than making some other decision.

On Alternative Accounts

Skeptical readers may worry about our interpretation of the behavior we observed. There are at least two alternative explanations those not familiar with the trust game might offer. The first is that we did not observe trust but rather altruism. The outcomes of other people matter, and people are willing to sacrifice some of their own material well-being to increase that of another person, even a stranger (Batson, 2011; Oord, 2007). In economic terms, our participants might have been motivated to increase efficiency; that is, to enlarge the pie of wealth and resources to the greatest possible extent, no matter who benefits (Hardin, 1968). Or, they may have been motivated to "do no harm" to the financial prospects of another person.

Our last two studies, as well as data elsewhere, speak against this alternative account. If people were motivated to give up their \$5 out of altruism or efficiency motives, they should have favored the coin flip game at least as much as the trust game. However, people were quite averse to giving the \$5 to Person B when that person would flip a coin to determine the fate of the \$20, despite such behavior being just as altruistic and efficient. Data from other studies shows this same pattern: Adding payoffs to other people beyond the self in a coin flip fails to inspire more gambling among participants (Bohnet & Zeckhauser, 2004; Schlösser, Fetchenhauer, & Dunning, 2013; Schlösser, Mensching, Dunning, & Fetchenhauer, 2014). Further, people are not influenced in the trust game by whether their interaction partner is poor, and thus needs the money, or is rich (Brühlhart & Usunier, 2012).

Readers may also speculate that our results are biased because participants made decisions only about small amounts of money

(e.g., \$5/\$20) that the experimenter provided to them. By this logic, people are more likely to gamble with "house money" of little consequence, so not much trust would be required. Past research, however, suggests that this concern is not valid. First, participants may be quite willing to hand their \$5 to a stranger, but if the stake is serious many more respondents refuse to gamble it in a lottery (Fetchenhauer & Dunning, 2012). Second, when we require people to bring their own money to the experiment, trust rates remain undiminished (Schlösser, Mensching, et al., 2014). Finally, economic researchers have varied the size of stake involved in trust games. Regardless of whether stakes are very high (in games played out in poor countries where the initial endowment was the income equivalent of \$1,693 in the United States in 2005) or very low (as in online studies), patterns of behavior remain largely unchanged (Amir, Rand, & Gal, 2012; Johansson-Stenman, Mahmud, & Martinsson, 2005).

Possible Extensions to Other Social Dilemmas

Although work on the trust game in psychology might be construed as nascent, the discipline has a long history of studying similar behavior in social dilemmas: mixed motive games, such as the prisoner's dilemma and public goods games, that feature choices between material self-interest versus acting for the benefit of all (for reviews, see Camerer, 2003; Parks, Joireman, & Van Lange, 2013).

Are there any lessons from the findings here that can be transferred to work on these other social dilemmas? After all, one might assert that trust and cooperation, for example, in the prisoner dilemma game, are the same construct, in that each involves a choice about making oneself vulnerable to exploitation by another. We would be cautious about such an assertion, however. There are many details between the trust game and other social dilemmas that differ, and these differences might activate very different psychological reactions. After all, other researchers have shown that seemingly trivial changes in the choice architecture of an economic game can have a powerful impact on whether the typical person acts in pro-social or selfish way (Bardsley, 2008; List, 2007).

For example, one major difference between the trust game and other social dilemmas is that the roles played by the players in the trust game are asymmetric: Person A must decide to give up the \$5 to allow Person B a decision that carries any consequence. As such, Person A has power over whether Person B can play meaningfully at all. Indeed, Person A really has the power to exit the game by merely keeping the money. The prisoner's dilemma differs in that both players face the same decision. In the traditional paradigm, they cannot exit the game. They must make a choice.

These differences might explain an apparent contradiction between our conclusions about the trust game and current thinking about games involving cooperation. We found that expectations about reward played a significant but hardly comprehensive role in people's decisions to trust, but work on other social dilemmas has found that social expectations play a large and seemingly central role (e.g., Balliet & Van Lange, 2013a). Why might that be?

We can furnish two relevant speculations. First, participants in the prisoner's dilemma, facing the exact same choice, may form expectations about their playing partner they feel are more valid and informative than those formed by participants in the trust.

Second, much work in social dilemmas has demonstrated that social norms are in operation by showing that people reward cooperators and punish the selfish. These carrots and sticks both work to increase cooperation (Balliet, Mulder, & Van Lange, 2011; Balliet & Van Lange, 2013b; Henrich et al., 2010). If behavior in social dilemmas is shaped more by norms that are social in nature, then social variables, such as expectation, might be more salient to participants. That is, whether or not people cooperate depends more on whether they think a social norm is in place, driving both their expectations and their behavior. Trust, driven more by private injunctions (Bicchieri et al., 2011), those considerations about others fail to rise to the same level of importance.

However, these contradictions about the role played by social expectation may be more apparent than real. We found that social expectations do matter but that they were not the only psychological mechanism in play. Other social dilemmas might also be influenced by moral norms (such as respect) or other mechanisms that have simply not yet been placed under empirical scrutiny. Thus, one can argue the possibility that both trust and other social dilemma behavior, once all the data are in, might turn out to be largely responsive to the same set of psychological dynamics (social expectation, moral norms) to similar degrees.

In addition, of greater import to our argument is the ability of norms to explain trust behavior beyond a baseline of economic rationality. We found that that moral norm appears to lift people's trust rates beyond a level anticipated by expectation and risk tolerance. To be sure, social expectation did matter, in that people tended to trust more when they were more optimistic about their peers' largesse, but those expectations could not explain why trust exceeded pure economic considerations.

At the end of the day, we might find a similar pattern in other social dilemmas: Social expectations may moderate who cooperates, for example, in the prisoner dilemma game. However, one key feature of that game is that it is never economically rational to cooperate. At every level of expectation, an economic analysis suggests that people should defect, and so the overall cooperation rate should be zero (Camerer, 2003). What, then, may account for positive cooperation rates? One possibility might be the presence of norms, personal or social, such as we have found here. Indeed, several researchers have posited the influence of a cooperative norm (e.g., Balliet et al., 2011) or investigated normative influence broadly (e.g., Bettenhausen & Murnighan, 1991), but, to our knowledge, no work has yet identified the specific boundaries of that norm or whether it works via social or personal avenues (Biel & Thøgersen, 2007, p. 105). Thus, again, data are needed to empirically test this possibility. Until then, it is still an open question whether trust operates via a different or roughly the same set of psychological mechanisms as other social dilemmas.

Final Words: Implications for Relationships Research

In summary, this research suggests that trust at zero acquaintance might be influenced more by internalized norms than by any consequentialist calculation about outcomes. This perspective leads to two general thoughts about future research and topics that may be worthy of serious consideration by scholars of interpersonal relationships.

First, one might argue that trust among strangers is not worthy of study because such transitory and ephemeral relationships are of

little to no importance. One need not worry about trusting the barista at one's favorite coffee shop, because any interaction one might have with that individual is trivial and routine. We think such an outlook, however, would be mistaken. As recent social theorists have pointed out, cultures often flourish only to the extent that people can come to rely on the benevolent actions of a whole host of strangers (Fehr et al., 2002; Fukuyama, 1995; Henrich, 2006; Henrich et al., 2010). Imagine a world in which one instead trusted only one's family and close acquaintances each day for food. Thus, determining how cultures develop and then maintain cooperative stances among strangers is a central question for how people create interpersonal social orders that are somewhat different from nasty, brutish, and short.

Second, trust in more established relationships has typically been defined and investigated in terms of beliefs and expectations within the relationship (e.g., Rousseau et al., 1998; Simpson, 2007); that is, whether people believe the other person will be responsive to and considerate of their personal welfare. This is a consequentialist perspective. One might raise the question of whether there is room for nonconsequentialist, norm-driven dynamics for trust in established relationships as well. One could ask, more specifically, if norm-driven trust might be an important catalyst for the development of established relationships (Lawler, 2001; Lawler, Thye, & Yoon, 2008). In the beginning of a relationship, the normative influences demonstrated here at zero acquaintance may be crucial to establish the dynamics of a trusting partnership; norms may be let to slip after partners get past the initial stage of potential nontrust. Once some small amount of trust is established, it produces "reciprocal reassurance" within the dyad and becomes self-perpetuating and self-enhancing (Holmes & Rempel, 1989). However, without norms to enable initial trust interactions, it is possible that many relationships (and exchanges) would have a more difficult time lifting off the ground, if they ever would.

References

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Akerlof, G., & Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, *115*, 715–753. doi:10.1162/003355300554881
- Amir, O., Rand, D., & Gal, Y. K. (2012). Economic games on the internet: The effect of \$1 stakes. *PLoS ONE*, *7*, e31461. doi:10.1371/journal.pone.0031461
- Anderson, J. E., & Dunning, D. (2014). *Behavioral norms: Variants and their identification*. Manuscript submitted for publication.
- Anderson, J. E., & Dunning, D. (n.d.). [Layperson ratings of labels to apply to the trust game]. Unpublished data, Cornell University.
- Andreoni, J. (1989). Giving with impure altruism: Applications to charity and Ricardian equivalence. *Journal of Political Economy*, *97*, 1447–1458. doi:10.1086/261662
- Ashraf, N., Bohnet, I., & Plankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, *9*, 193–208. doi:10.1007/s10683-006-9122-4
- Bacharach, M., & Gambetta, D. (2001). Trust in signs. In K. Cook (Ed.), *Trust in society* (Vol. 2, pp. 148–184). New York, NY: Russell Sage Foundation.
- Baldwin, M. W., Fehr, B., Keedian, E., Seidel, M., & Thompson, D. W. (1993). An exploration of the relational schemata underlying attachment styles: Self-report and lexical decision approaches. *Personality and Social Psychology Bulletin*, *19*, 746–754. doi:10.1177/0146167293196010

- Balliet, D., Mulder, L. B., & Van Lange, P. A. M. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin, 137*, 594–615. doi:10.1037/a0023489
- Balliet, D., & Van Lange, P. A. M. (2013a). Trust, conflict, and cooperation: A meta-analysis. *Psychological Bulletin, 139*, 1090–1112. doi:10.1037/a0030939
- Balliet, D., & Van Lange, P. A. M. (2013b). Trust, punishment and cooperation across 18 societies: A meta-analysis. *Perspectives on Psychological Science, 8*, 363–379. doi:10.1177/1745691613488533
- Barber, B. (1983). *The logic and limits of trust*. New Brunswick, NJ: Rutgers University Press.
- Bardsley, N. (2008). Dictator game giving: Altruism or artefact? *Experimental Economics, 11*, 122–133. doi:10.1007/s10683-007-9172-2
- Batson, C. D. (2011). *Altruism in humans*. New York, NY: Oxford University Press.
- Batson, C. D., & Moran, T. (1999). Empathy-induced altruism in a prisoner's dilemma. *European Journal of Social Psychology, 29*, 909–924. doi:10.1002/(SICI)1099-0992(199911)29:7<909::AID-EJSP965>3.0.CO;2-L
- 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. doi:10.1016/j.joep.2009.10.001
- 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
- Bettenhausen, K. L., & Murnighan, J. K. (1991). The development of an intragroup norm and the effects of interpersonal and structural changes. *Administrative Science Quarterly, 36*, 20–35. doi:10.2307/2393428
- Bicchieri, C. (2006). *The grammar of society*. New York, NY: Cambridge University Press.
- Bicchieri, C., Xiao, E., & Muldoon, R. (2011). Trustworthiness is a norm, but trusting is not. *Politics, Philosophy & Economics, 10*, 170–187. doi:10.1177/1470594X10387260
- Biel, A., & Thøgersen, J. (2007). Activation of social norms in social dilemmas: A review of the evidence and reflections on the implications for environmental behaviour. *Journal of Economic Psychology, 28*, 93–112. doi:10.1016/j.joep.2006.03.003
- Bodner, R., & Prelec, D. (2002). Self-signaling and diagnostic utility in everyday decision making. In I. Brocas & J. Carillo (Eds.), *Collected essays in psychology and economics* (pp. 1–22). New York, NY: Oxford University Press.
- 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
- Bolle, F. (1998). Rewarding trust: An experimental study. *Theory and Decision, 45*, 83–98. doi:10.1023/A:1004930101743
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and cooperation. *American Economic Review, 90*, 166–193. doi:10.1257/aer.90.1.166
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429–444. doi:10.1111/0022-4537.00126
- Bronfenbrenner, U. (1970). *Two worlds of childhood: US and USSR*. New York, NY: Simon & Schuster.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge, United Kingdom: Cambridge University Press.
- Brühlhart, M., & Usunier, J.-C. (2012). Does the trust game measure trust? *Economic Letters, 115*, 20–23. doi:10.1016/j.econlet.2011.11.039
- Burnham, T., McCabe, K., & Smith, V. L. (2000). Friend-or-foe intentionality priming in an extensive form trust game. *Journal of Economic Behavior & Organization, 43*, 57–73. doi:10.1016/S0167-2681(00)00108-6
- Camerer, C. F. (2003). *Behavioral game theory: Experiments on strategic interaction*. Princeton, NJ: Princeton University Press.
- Chudek, M., & Henrich, J. (2011). Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences, 15*, 218–226. doi:10.1016/j.tics.2011.03.003
- Cronk, L. (2007). The influence of cultural framing on play in the trust game: A Maasai example. *Evolution and Human Behavior, 28*, 352–358. doi:10.1016/j.evolhumbehav.2007.05.006
- Crosno, R., & Buchan, N. (1999). Gender and culture: International experimental evidence from trust games. *American Economic Review, 89*, 386–391. doi:10.1257/aer.89.2.386
- DePaulo, B. M., & Bell, K. L. (1996). Truth and investment: Lies are told to those who care. *Journal of Personality and Social Psychology, 71*, 703–716. doi:10.1037/0022-3514.71.4.703
- Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution, 2*, 265–279. doi:10.1177/002200275800200401
- Deutsch, M. (1973). *The resolution of conflict: Constructive and destructive processes*. New Haven, CT: Yale University Press.
- Dovidio, J. F., Gaertner, S. L., John, M., Halabi, S., Saguy, T., Pearson, A. R., & Riek, B. M. (2008). Majority and minority perspectives in intergroup relations: The role of contact, group representations, threat, and trust in intergroup conflict and reconciliation. In A. Nadler, T. E. Malloy, & J. D. Fisher (Eds.), *The social psychology of intergroup reconciliation* (pp. 227–253). New York, NY: Oxford University Press.
- Dubreuil, B., & Grégoire, J.-F. (2013). Are moral norms distinct from social norms? A critical assessment of Jon Elster and Cristina Bicchieri. *Theory and Decision, 75*, 137–152. doi:10.1007/s11238-012-9342-3
- Dunn, E. W., Aknin, L. B., & Norton, M. I. (2008, March 21). Spending money on others promotes happiness. *Science, 319*, 1687–1688. doi:10.1126/science.1150952
- Dunning, D., & Fetchenhauer, D. (2010a). Trust as an expressive rather than an instrumental act. *Advances in Group Processes, 27*, 97–127.
- Dunning, D., & Fetchenhauer, D. (2010b). Understanding the psychology of trust. In D. Dunning (Ed.), *Social motivation* (pp. 147–170). New York, NY: Psychology Press.
- Dunning, D., & Fetchenhauer, D. (2013). Behavioral influences in the present tense: On expressive versus instrumental action. *Perspectives on Psychological Science, 8*, 142–145. doi:10.1177/1745691612474319
- Dunning, D., Fetchenhauer, D., & Schlösser, T. (2012). Trust as a social and emotional act: Noneconomic considerations in trust behavior. *Journal of Economic Psychology, 33*, 686–694. doi:10.1016/j.joep.2011.09.005
- Eckel, C., & Wilson, R. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization, 55*, 447–465. doi:10.1016/j.jebo.2003.11.003
- Elster, J. (2007). *Explaining social behavior: More nuts and bolts for the social sciences*. Cambridge, United Kingdom: Cambridge University Press.
- Elster, J. (2009). Social norms and the explanation of behavior. In P. Hedström & P. Bearman (Eds.), *The Oxford handbook of analytical sociology* (pp. 195–217). Oxford, United Kingdom: Oxford University Press.
- Evans, A. M., & Krueger, J. I. (2009). The psychology (and economics) of trust. *Social and Personality Psychology Compass, 3*, 1003–1017. doi:10.1111/j.1751-9004.2009.00232.x
- Evans, A. M., & Krueger, J. I. (2011). Elements of trust: Risk and perspective-taking. *Journal of Experimental Social Psychology, 47*, 171–177. doi:10.1016/j.jesp.2010.08.007
- Evans, A. M., & Reville, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality, 42*, 1585–1593. doi:10.1016/j.jrp.2008.07.011
- Fehr, E. (2004, November 25). Don't lose your reputation. *Nature, 432*, 449–450. doi:10.1038/432449a

- Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of norms. *Human Nature, 13*, 1–25. doi:10.1007/s12110-002-1012-7
- Fehr, E., Fischbacher, U., von Rosenblatt, B., Schupp, J., & Wagner, G. (2002). A nation-wide laboratory: Examining trust and trustworthiness of integrating experiments in representative surveys. *Schmollers Jahrbuch, 122*, 519–542.
- 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 in the trustworthiness of others. *Psychological Science, 21*, 189–193. doi:10.1177/0956797609358586
- 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 & Organization, 81*, 534–541. doi:10.1016/j.jebo.2011.07.017
- Fetchenhauer, D., & Van der Vegt, G. (2001). Honesty, trust and economic growth: A cross-cultural comparison of Western industrialized countries. *Zeitschrift für Sozialpsychologie, 32*, 189–200. doi:10.1024/0044-3514.32.3.189
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York, NY: Psychology Press.
- Fiske, S. T. (2003). Five core social motives, plus or minus five. In S. J. Spencer, S. Fein, M. Zanna, & J. Olson (Eds.), *Motivated social perception: The Ontario Symposium* (Vol. 9, pp. 233–246). Mahwah, NJ: Erlbaum.
- Fraley, R. C., & Shaver, P. R. (2000). Adult romantic attachment: Theoretical developments, emerging controversies, and unanswered questions. *Review of General Psychology, 4*, 132–154. doi:10.1037/1089-2680.4.2.132
- Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. New York, NY: Free Press.
- Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics, 115*, 811–846. doi:10.1162/003355300554926
- Goffman, E. (1958). *The presentation of self in everyday life*. New York, NY: Random House.
- Hardin, G. (1968, December 13). The tragedy of the commons. *Science, 162*, 1243–1248. doi:10.1126/science.162.3859.1243
- Hartmann, N. (1932). *Ethics* (Vol. 2). New York, NY: Macmillan.
- Haselhuhn, M. P., & Mellers, B. A. (2005). Emotions and cooperation in economic games. *Cognitive Brain Research, 23*, 24–33. doi:10.1016/j.cogbrainres.2005.01.005
- Henrich, J. (2006, April 7). Cooperation, punishment, and the evolution of human institutions. *Science, 312*, 60–61. doi:10.1126/science.1126398
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., . . . Ziker, J. (2010, March 19). Markets, religion, community size, and the evolution and punishment. *Science, 327*, 1480–1484. doi:10.1126/science.1182238
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319–340. doi:10.1037/0033-295X.94.3.319
- Higgins, E. T. (1989). Continuities and discontinuities in self-regulatory and self-evaluative processes: A developmental theory relating self and affect. *Journal of Personality, 57*, 407–444. doi:10.1111/j.1467-6494.1989.tb00488.x
- Higgins, E. T., Bond, R. N., Klein, R., & Strauman, T. (1986). Self-discrepancies and emotional vulnerability: How magnitude, accessibility, and type of discrepancy influence affect. *Journal of Personality and Social Psychology, 51*, 5–15. doi:10.1037/0022-3514.51.1.5
- Higgins, E. T., Shah, J., & Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology, 72*, 515–525. doi:10.1037/0022-3514.72.3.515
- Hobbes, T. (1997). *The Leviathan*. New York, NY: Touchstone Press. (Original work published 1660)
- Holmes, J. G., & Rempel, J. K. (1989). Trust in close relationships. In C. Hendrick (Ed.), *Close relationships* (pp. 187–220). Newbury Park, CA: Sage.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York, NY: Wiley.
- 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
- Johnson, N. D., & Mislin, A. (2012). How much should we trust the World Values Survey trust question? *Economic Letters, 116*, 210–212. doi:10.1016/j.econlet.2012.02.010
- Kanagaretnam, K., Mestelman, S., Nainar, K., & Shehata, M. (2009). The impact of social value orientation and risk attitudes on trust and reciprocity. *Journal of Economic Psychology, 30*, 368–380. doi:10.1016/j.joep.2008.12.003
- Kay, A. C., & Ross, L. (2003). The perceptual push: The interplay of implicit cues and explicit situational construals on behavioral intentions in the Prisoner's Dilemma. *Journal of Experimental Social Psychology, 39*, 634–643. doi:10.1016/S0022-1031(03)00057-X
- Kelley, H. H., & Thibault, J. W. (1978). *Interpersonal relationships: A theory of interdependence*. New York, NY: Wiley.
- Klapwijk, A., & Van Lange, P. A. M. (2009). Promoting cooperation and trust in “noisy” situations: The power of generosity. *Journal of Personality and Social Psychology, 96*, 83–103. doi:10.1037/a0012823
- Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics, 112*, 1251–1288. doi:10.1162/003355300555475
- Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005, June 2). Oxytocin increases trust in humans. *Nature, 435*, 673–676. doi:10.1038/nature03701
- Kramer, R. M. (1998). Paranoid cognition in social systems: Thinking and acting in the shadow of doubt. *Personality and Social Psychology Review, 2*, 251–275. doi:10.1207/s15327957pspr0204_3
- Kramer, R. M., & Carnevale, P. J. (2001). Trust and intergroup negotiation. In R. Brown & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Vol. 4. Intergroup relations* (pp. 431–450). Oxford, United Kingdom: Blackwell.
- Kreps, D. M. (1990). Corporate culture and economic theory. In J. Alt & K. Shepsel (Eds.), *Perspectives on positive political economy* (pp. 90–143). New York, NY: Cambridge University Press.
- Kugler, T., Connolly, T., & Ordóñez, L. D. (2012). Emotion, decision, and risk: Betting on gambles versus betting on people. *Journal of Behavioral Decision Making, 25*, 123–134. doi:10.1002/bdm.724
- Lacetera, N., & Macis, M. (2010). Self-image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior & Organization, 31*, 738–748.
- Lawler, E. J. (2001). An affect theory of social exchange. *American Journal of Sociology, 107*, 321–352. doi:10.1086/324071
- Lawler, E. J., Thye, S. R., & Yoon, J. (2008). Social exchange and micro social order. *American Sociological Review, 73*, 519–542. doi:10.1177/000312240807300401
- Lazzarini, S. G., Madalozzo, R., Artes, R., & Siqueira, J. O. (2005). Measuring trust: An experiment in Brazil. *Revista de Economia Aplicada, 9*, 153–169.
- Lieberman, V., Samuels, S. M., & Ross, L. (2004). The name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. *Personality and Social Psychology Bulletin, 30*, 1175–1185. doi:10.1177/0146167204264004

- Lind, E. A. (1988). *The social psychology of procedural justice*. New York, NY: Plenum Press.
- List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, *115*, 482–493. doi:10.1086/519249
- Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family: Parent–child interaction. In P. H. Mussen & E. M. Hetherington (Eds.), *Handbook of child psychology: Vol. 4. Socialization, personality, and social development* (pp. 1–101). New York, NY: Wiley.
- Machiavelli, N. (1515/2003). *The prince* (R. Goodwin, Trans.). Wellesley, MA: Dante University Press.
- Mellers, B. A. (2000). Choice and the relative pleasure of consequences. *Psychological Bulletin*, *126*, 910–924. doi:10.1037/0033-2909.126.6.910
- Mellers, B. A., Ritov, I., & Schwartz, A. (1999). Emotion-based choice. *Journal of Experimental Psychology: General*, *128*, 332–345. doi:10.1037/0096-3445.128.3.332
- Mikulincer, M. (1998). Attachment working models and the sense of trust: An exploration of interaction goals and affect regulation. *Journal of Personality and Social Psychology*, *74*, 1209–1224. doi:10.1037/0022-3514.74.5.1209
- Milinski, M., Semmann, D., & Krambeck, H. J. (2002, January 24). Reputation helps solve the “tragedy of the commons.” *Nature*, *415*, 424–426. doi:10.1038/415424a
- Miller, D. T. (2006). *An invitation to social psychology: Expressing and censoring the self*. Belmont, CA: Wadsworth-Thomson.
- Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*, 33–43. doi:10.1037/0022-3514.81.1.33
- Murray, S. L., Holmes, J. G., & Collins, N. L. (2006). Optimizing assurance: The risk regulation system in relationships. *Psychological Bulletin*, *132*, 641–666. doi:10.1037/0033-2909.132.5.641
- Naef, M., & Schupp, J. (2009). *Measuring trust: Experiments and surveys in contrast and combination*. Berlin, Germany: DIW Berlin.
- Oishi, S., Kesebir, S., & Diener, E. (2011). Income inequality and happiness. *Psychological Science*, *22*, 1095–1100. doi:10.1177/0956797611417262
- Olivola, C. Y., & Todorov, A. (2010). Fooled by first impressions? Re-examining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*, 315–324. doi:10.1016/j.jesp.2009.12.002
- Oord, T. (2007). *The altruism reader*. Philadelphia, PA: Templeton Foundation Press.
- Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, *9*, 128–133. doi:10.1037/a0014520
- Parks, C. D., Joireman, J., & Van Lange, P. A. M. (2013). Cooperation, trust, and antagonism: How public goods are promoted. *Psychological Science in the Public Interest*, *14*, 119–165. doi:10.1177/1529100612474436
- Pillutla, M. M., Malhotra, D., & Murnighan, J. K. (2003). Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology*, *39*, 448–455. doi:10.1016/S0022-1031(03)00015-5
- Posten, A.-C., Ockenfels, A., & Mussweiler, T. (2013). How activating cognitive content shapes trust: A subliminal framing study. *Journal of Economic Psychology*. Advance online publication. doi:10.1016/j.joep.2013.04.002
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891. doi:10.3758/BRM.40.3.879
- Pruitt, D. G., & Rubin, J. Z. (1986). *Social conflict: Escalation, stalemate, and settlement*. New York, NY: McGraw-Hill.
- Putnam, R. D. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.
- Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, *49*, 95–112. doi:10.1037/0022-3514.49.1.95
- Risen, J. L., & Gilovich, T. (2007). Target and observer differences in the acceptance of questionable apologies. *Journal of Personality and Social Psychology*, *92*, 418–433. doi:10.1037/0022-3514.92.3.418
- Roney, C. J. R., & Sorrentino, R. M. (1995). Reducing self-discrepancies or maintaining self-congruence? Uncertainty orientation, self-regulation, and performance. *Journal of Personality and Social Psychology*, *68*, 485–497. doi:10.1037/0022-3514.68.3.485
- Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, *35*, 651–665. doi:10.1111/j.1467-6494.1967.tb01454.x
- Rotter, J. B. (1971). Generalized expectancies of interpersonal trust. *American Psychologist*, *26*, 443–452. doi:10.1037/h0031464
- Rousseau, D. M., Sitkin, S. B., 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
- Schlösser, T., Fetchenhauer, D., & Dunning, D. (2013). *Trust against all odds? Emotional dynamics in trust behavior*. Unpublished manuscript, University of Cologne, Cologne, Germany.
- Schlösser, T. M., Mensching, O., Dunning, D., & Fetchenhauer, D. (2014). *Why do people trust? A multi-perspective approach to explain the phenomenon of trust*. Unpublished manuscript, University of Cologne, Cologne, Germany.
- Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The “big three” of morality (autonomy, community, divinity), and the “big three” explanations of suffering. In P. Rozin & A. Brandt (Eds.), *Morality and health* (pp. 119–169). New York, NY: Routledge.
- Simpson, J. A. (2007). Foundations of interpersonal trust. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 587–607). New York, NY: Guilford Press.
- Snijders, C., & Keren, G. (2001). Do you trust? Whom do you trust? When do you trust? *Advances in Group Processes*, *18*, 129–160. doi:10.1016/S0882-6145(01)18006-9
- Sullivan, J. L., & Transue, J. E. (1999). The psychological underpinnings of democracy: A selective review of research on political tolerance, interpersonal trust, and social capital. *Annual Review of Psychology*, *50*, 625–650. doi:10.1146/annurev.psych.50.1.625
- Swim, J. K., & Hyers, L. L. (1999). Excuse me—What did you just say? Women’s public and private responses to sexist remarks. *Journal of Experimental Social Psychology*, *35*, 68–88. doi:10.1006/jesp.1998.1370
- Turiel, E. (1983). *The development of social knowledge*. Cambridge, United Kingdom: Cambridge University Press.
- Uslaner, E. M. (2002). *The moral foundations of trust*. Cambridge, United Kingdom: Cambridge University Press.
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100-ms exposure to a face. *Psychological Science*, *17*, 592–598. doi:10.1111/j.1467-9280.2006.01750.x
- Wilson, R., & Eckel, C. C. (2011). Trust and social exchange. In J. N. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *The handbook of experimental political science* (pp. 243–257). Boston, MA: Cambridge University Press.
- Woodzicka, J. A., & LaFrance, M. (2001). Real versus imagined gender harassment. *Journal of Social Issues*, *57*, 15–30. doi:10.1111/0022-4537.00199

Received October 10, 2013

Revision received March 10, 2014

Accepted March 13, 2014 ■