

## Trust: The Essential Social Dilemma

Joachim I. Krueger<sup>1</sup> and Anthony M. Evans<sup>2</sup>

<sup>1</sup>Brown University <sup>2</sup>Tilburg University

### Keywords

Trust, power, freedom, social dilemmas decision-making

*A social organism of any sort whatever, large or small, is what it is because each member proceeds to his own duty with a trust that the other members will simultaneously do theirs.*

William James (2007/1897, p. 24)

*Trust in me, just trust in me.  
Shut your eyes and trust in me.*  
Kaa to Mowgli in *The Jungle Book*

Decisions of whether to trust others pervade human affairs. These decisions are unavoidable in a social species whose members are able to take the perspective of others. People appreciate that these others are motivated by self-interest, but also know that self-interest must bend to the need to get along with others. Online commerce is an example of a trust-based interaction among strangers. Even individuals who know each other well can never take the other's trustworthiness for granted. Betrayal may be improbable, but it is not impossible. In this article, we discuss trust in the context of two other social motives, freedom and power. Having shown that the dilemma (see glossary) of trust has no easy solution, we review four strategies to mitigate it.

### Freedom

A science-fiction writer can imagine a world in which human behavior is perfectly predictable (Dick, 2002/1956). In the real world, behavior is predictable only within limits and often only in the aggregate. Psychological science seeks to push back these limits. What if science triumphed? We may prefer the behavior of others to be predictable, but if we got our wish, our own behavior would also be predictable by others. As a result, we would no



**Fig. 1.** Shut your eyes and trust in me.

longer be free.<sup>1</sup> Freedom is the ability to behave in unexpected ways. We can use this freedom to our advantage, but also to the advantage of others. Without the freedom to be unpredictable, we cannot deliver pleasant surprises. At the same time, behavioral freedom (see glossary) implies the freedom to betray, cheat, and deceive (Miller, 1996; Nozick, 2001). We cannot have one without the other. Where the ability to predict the behavior of others is limited, trust and suspicion become important.

All humans must trust – at least some of the time. A person who is always suspicious, that is, someo-

<sup>1</sup> The freedom to choose does not require metaphysical free will. All human behavior is embedded in causal networks (and perhaps irreducible elements of chance). The experience of being free is precious, if illusory.

### Corresponding author:

Joachim I. Krueger  
Department of Cognitive, Linguistic & Psychological Sciences  
Brown University  
Box 1853 190 Thayer St. Providence, RI 02912  
E-mail: [Joachim\\_Krueger@Brown.edu](mailto:Joachim_Krueger@Brown.edu)

ne who always expects deceit and betrayal believes that the behavior of others is predictable, but such a person cannot flourish in a social species. Likewise, someone who always trusts cannot flourish because such a person will eventually be exploited. The freedom to bluff and betray others some of the time is balanced by the freedom to trust or be suspicious.

Psychologists and behavioral economists know that trust is a dilemma. According to a widely accepted definition, trust is “a psychological state comprising the intention to accept vulnerability based upon the positive expectations of the intentions or behavior of another” (Rousseau, Sitkin, Burt, & Camerer, 1998, p. 395). People who trust accept the possibility that they might be betrayed inasmuch as they think that this possibility is small. The noted sociologist Niklas Luhmann (2000) regarded trust as a “riskante Vorleistung”, or risky forward payment. He distinguished trust from hope, which is a sentiment that remains when the person has no option to decline the risky behavior.

Trust is not simply a matter of computing the expected value of a potential interaction and act accordingly. The reason trust is a dilemma is that the probability of another person’s betrayal is difficult to estimate, and the source of this difficulty is freedom. Someone who wants to cheat others will want to hide her intentions. In contrast, someone who has no intention of cheating has no incentive to pretend otherwise. Would-be trustors can gather all the available cues, estimate the probability that their trust will be reciprocated – and adjust downward. The trouble is that no cues are left to indicate how large that adjustment should be.

## Power

While people value trust and trustworthiness, they also value power. Interpersonal power (see glossary) refers to the ability to control resources, or more broadly, to reward and punish. The writer of Job (1:21) is resigned to conclude that “The Lord gave and the Lord hath taken away.” When Job becomes the victim of arbitrary and thus absolute power, he can no longer trust that his righteous ways will be rewarded. He learns that trust and power are intertwined. If the behavior of individuals perfectly predicted the responses of the powerful (gods, parents, bosses), that power would be negated at the same time. The presumably powerless individuals, with their freedom to experiment, would gain control.<sup>2</sup> Conversely, if the powerful fully executed

their freedom to deliver rewards and punishments at will, their choices would be unrelated to the behavior they presumably address, and hence there could be no trust (cf. Feuerbach, 1989/1830). In ordinary social life, such extremes are rare. Typically, there is a zone in which trust and trustworthiness are correlated. People who trust can expect that their acceptance of risk and vulnerability will be noticed and rewarded probabilistically. Conversely, people who are being trusted, tend to respond to the normative obligation to reward trust (Gouldner, 1960), but they are rarely completely reliable. Those who want to protect their freedom or their power must on occasion frustrate those who trust them.

Bertrand Russell (1938) proposed that all social behavior could be understood through the lens of power. To Russell, power was to social relations what energy is to physics. Other social scientists emphasize trust over power. Nobel Laureate of Economics Kenneth Arrow (1974, p. 23) asserted that trust is “a lubricant for social systems,” and psychologist Julian Rotter (1970, p. 443) saw trust everywhere. “The entire fabric of our day-to-day living, of our social world, rests on trust – buying gasoline, paying taxes, going to the dentist, flying to a convention – almost all our decisions involve trusting someone else.”

We submit that neither trust nor power can be understood in isolation; if either one were to completely dominate, life would be grim. If everyone always trusted everyone else, society might be efficient but boring. The concept of trust would lose its meaning because there would be no risk left. However, a society of trustors would be vulnerable to the invasion by genetic or social mutants ready to exploit and displace them unless the trustors learned to be strategic. Likewise, if everyone related to others only in terms of power, society would be a Hobbesian war of all against all. The challenge is to find a balance between trust and power and to find the sweet spot that maximizes individual and social well-being (see also Krueger, 2013b, for an informal discussion).

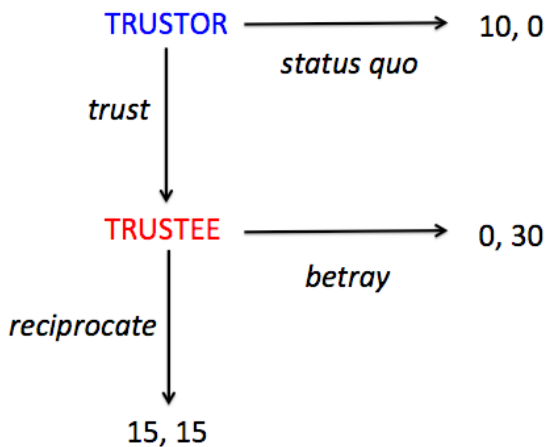
## The Game

In experimental psychology and behavioral economics, trust and its counterpart of trustworthiness are studied in a paradigm borrowed from game theory (see glossary). In the simplest version of this game, a player, the trustor, can choose to keep a small amount of money (e.g., €10), or transfer it to another player, the trustee. If she transfers the money, the amount is tripled (by the experimenter playing the role of the invisible hand). If the trustee receives €30, she

<sup>2</sup> In its weak sense, control is merely predictive. John might know that Paul will slap him every Tuesday morning at 9:00 a.m. In its strong sense, control is causal. If intercessory

prayer were perfectly successful, god would be reduced to an instrument.

can either split the amount with the trustor or keep it all. Figure 1 shows the extensive form of the game. Research shows that rates of trust and trustworthiness are intermediate and that there is a great deal of variation over studies (Evans & Krueger, 2009).



**Fig. II.** The extensive form of the trust game. The values to the left and right of the comma respectively go to the trustor and the trustee.

This brief summary allows the following conclusions: [1] The empirical evidence for trust and trustworthiness refutes the strict game-theoretic view that neither should occur at all.<sup>3</sup> [2] Being less than perfect, trust fails to create maximum wealth, thereby leading to calls for greater trust. [3] Because it is moderate instead of complete, trust preserves its conceptual integrity (see above) and a state of balance with power. Being only partly predictable, players assert their behavioral freedom and the power to say no.

When the issue of power is raised, it typically concerns the trustee. Once trust has occurred, there is no more uncertainty for the trustee, who now gets to play a dictator game (Hoffman, McCabe, & Smith, 1996). She has control over the resource (€30) and can divide it at will, with the trustor at her mercy. Her decision is simply a matter of finding the balance between self-interest and sensitivity to the norm of reciprocity. This does not mean, however, that the role of trustor is without power. At the outset, the trustor controls the resource (€10). At this stage, the trustee's well-being is up to him. Still, power is transient in a one-shot game. Once it is used, it is gone.

Most people value freedom and power, but they also want more trustworthiness from others. What about trust? Unlike freedom or power, trust is not obtained by overcoming resistance; one can just exhibit more of it by virtue of one's freedom

<sup>3</sup> Not trusting (and hence not receiving anything back) is a Nash equilibrium. Rational players do not trust because they (presumably) know that there is no reason to reciprocate trust.

and power. If given a choice, would a person rather play the role of trustor or trustee? We surveyed students in the laboratory and classroom and found no clear preference. It appears that in the game, trust is balanced with freedom and power.

## Beyond Balance

Trust situations demand a balance among competing strategic forces and motives, but this balance is fragile. Societies with higher levels of trust are more cohesive and productive than societies with low levels of trusts, and the effect appears to be causal (Putterman, 2012). In "The Price of Inequality," Nobel Laureate of Economics Joseph Stiglitz (2012) writes that trust depends on the perception of fairness, and that trusting individuals work harder and contribute more to society; in other words, they create the social capital that "is the glue that holds society together" (p. 122). Since the Great Recession of 2008, the industrialized world has seen trust erode. Stiglitz reports that Gallup and New York Times surveys respectively show steep losses of trust in business and government. Within the business sector, the banking industry has suffered, perhaps deservedly, the most. This is tragic because banking used to be the quintessentially trust-based industry.

Business writers have discovered the importance of trust, but, like most academic researchers, they construe trust as being anadromous, or something that flows upstream in social hierarchies (see Horsager, 2011, for a broader view). There is little work addressing leaders' trust in their employees. This is a grave omission because social relations – power differentials notwithstanding – are bidirectional. A hypothesis worth considering is that trusting leaders are highly productive because they avoid micro-management. Trusting leaders may set in motion a positive feedback loop that generates new social capital (see McGregor, 1960, and Akerlof & Kranton, 2005, respectively for a classic and a contemporary theoretical treatment). The idea was Lao-tzu's: "He who does not trust enough, is not to be trusted" (Tao Te Ching, verse 17; see also Murnighan, 2012, for a Neo-Taoist view on leadership).

Research on repeated games, where the same partners interact numerous times, shows that trust begets trust in a virtuous cycle of reciprocity (Wedekind & Milinski, 1996). Those who do not trust or stop interacting after one disappointment, will underestimate the trustworthiness of individuals or groups. Those who are willing to trust will over time form more accurate perceptions of trustworthiness, even if at first they overestimate its prevalence (Denrell & Le Mens, 2012).

## Raising Trust

What can be done when trust falls below balanced levels? The way of the Tao is to initiate trust unilaterally, hoping for positive returns. This advice is sage but hard to follow if there is no other reason to motivate trust. We hereby consider four strategies to increase trust. None of them is foolproof, which reinforces the idea that the dilemma of trust cannot be conclusively solved without defining it away.

The first strategy is to stimulate social projection (see glossary), which can be accomplished by pointing out similarities among the interactants. Social projection is expressed by the expectation that others will act as oneself would (Krueger, 2013a). A trustor may feel that her own socially beneficial choice (trust) is diagnostic of the other's socially beneficial choice (reciprocity; see Krueger, DiDonato, & Freestone, 2012, for a formal treatment). A trustor can perform projection by mentally simulating what she would do if she were the trustee. If she believes she would reciprocate trust, she herself might trust. The risk is that once she takes the trustee's perspective, she might realize (or even overestimate) the allure of betrayal (Epley, Caruso, & Bazerman, 2006).

The second strategy is to increase the efficiency gain produced by trust. In the trust game, the tripling of the transferred money is a matter of convention. Trust increases with a larger multiplier and social efficiency is enhanced (Johnson & Mislin, 2011). On the one hand, trust may come easier with a multiplier of 5 rather than 3 because a lower probability of reciprocity is sufficient to make the expected value of trust larger than the expected value of distrust ( $p > .4$  vs.  $.67$ ). On the other hand, the trustor should realize that the temptation of betrayal has also increased, thus lowering the probability of reciprocity. Increasing the multiplier increases trust inasmuch as trustors are more focused on their own improved prospects than on the improved prospects of the trustee (Evans & Krueger, 2011).<sup>4</sup>

The third strategy is to make trust the default. In the standard trust game, trustors must take action to transfer money to the trustee. Doing nothing amounts to distrust. Defaults dominate behavior when decisions are difficult (Johnson et al., 2012). Hence, a lack of trust is, in part, a matter of the default, or status quo, bias. In a laboratory study, we found greater trust when trustors did not need to actively transfer money, but this effect occurred only when cognitive resources had been taxed by a pre-

vious task (Evans, Dillon, Goldin, & Krueger, 2011).

The fourth strategy is to signal trustworthiness with nonverbal behavior (see glossary) when the opportunity arises in face-to-face interaction (DeSteno, Breazeal et al., 2012). Individuals who come across as agreeable are trusted more (Todorov, 2008). In dynamic interaction, it pays (literally) to express agreeableness by mimicking the other person (Mad-dux, Mullen, & Galinsky, 2008). Finally, although we do not recommend trying this outside the laboratory, the administration of oxytocin (the "love hormone") increases trust (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005). Then again, claiming to be a doctor may be enough to (see picture above) .

## Conclusion

Trust is a dilemma that is not easily solved. Humans will continue to worry about it. As William James (1897) noted, people of goodwill are not cynical about the intention of others. A world in which trust is impossible or unnecessary would not be rewarding or even interesting. Still, humans must stay alert to the possibility of abuse, as Mowgli would have done well to distrust Kaa. When there is a crisis of trust, as in today's economic world, efforts to restore it should be encouraged – along with efforts to restore trustworthiness, that is. Meanwhile, psychological science continues to seek a better understanding of trust, how it works and how it is related to other social concepts.

## Glossary

**Behavioral freedom.** Freedom may be defined in metaphysical or political terms, but behavioral freedom refers to an organism's ability to vary its behavior in such a way that it becomes difficult to predict.

**Game theory.** A branch of applied mathematics and influential framework for the study of economics. Game theory assumes that individuals are rational, self-interested agents, who have a set of stable and well-defined preferences and who know this about one another. Game theory derives mathematically the choices of such individuals.

**Interpersonal power.** According to a narrow definition, interpersonal power resides in the control over resources valued by others. According to a broad definition, power involves the ability to influence the behavior of others. The distinction is academic because a person who has power in the strict sense, will use it to achieve results in the broad sense.

**Social dilemma.** Types of games characterized by a misalignment of individual and social preferences. What is best for the individual is not best for the group. The prisoner's dilemma, public-goods dilemmas, resource-depletion dilemmas are well-known examples. The trust

4. A practical problem is that this strategy requires an increase in transaction benefits that must come from somewhere. There is an element of question begging. The strategy purports to improve efficiency by improving efficiency.

game has the defining characteristics of a social dilemma.

**Social projection.** The perception of similarity between self and other as well as the processes that bring about this perception. Social projection is a special case of inductive reasoning, where a person estimates the unknown properties of general targets (a social category, a group, other people) on the basis of his or her own properties.

**Nonverbal behavior.** Any behavior that is not verbal, that is, the great majority of what people (and other animals) do. Among humans, most nonverbal communication flows through the face, with gaze being most significant. Pose communicates power, but not trust (worthiness).

## References

- Akerlof, G. A., & Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, *19*, 9-32.
- Arrow, K. (1974). *The limits of organization*. New York, NY: Norton & Company, Inc.
- Denrell, J., & Le Mens, G. (2012). Social judgments from adaptive samples. In J. I. Krueger (Ed.), *Social judgment and decision making* (pp. 151-169). New York, NY: Psychology Press.
- DeSteno, D., Breazeal, C., Frank, R. H., Pizarro, D., Baumann, J., Dickens, L., ... Lee, J. J. (2012). Detecting the trustworthiness of novel partners in economic exchange. *Psychological Science*, *23*, 1549-1556. doi: 10.1177/0956797612448793
- Dick, P. K. (2002/1956). *The minority report*. London, UK: Golancz.
- Epley, N., Caruso, E., & Bazerman, M. H. (2006). When perspective-taking increases taking: Reactive egoism in social interaction. *Journal of Personality and Social Psychology*, *91*, 872-889. doi: 10.1037/0022-3514.91.5.872
- Evans, A. M., Dillon, K. D., Goldin, G., & Krueger, J. I. (2011). Trust and self-control: The moderating role of the default. *Judgment and Decision Making*, *6*, 697-705.
- Evans, A. M., & Krueger, J. I. (2009). The psychology (and economics) of trust. *Social and Personality Psychology Compass: Intrapersonal Processes*, *3*, 1003-1017. doi:10.1111/j.1751-9004.2009.00232.x
- Evans, A. M., & Krueger, J. I. (2011). Elements of trust: Risk taking and expectation of reciprocity. *Journal of Experimental Social Psychology*, *47*, 171-177. doi:10.1016/j.jesp.2010.08.007
- Feuerbach, L. A. (1989). *The essence of Christianity [Das Wesen des Christentums]* (G. Eliot, Trans.). Buffalo, NY: Prometheus Books (Original work published 1830).
- Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, *25*, 161-178.
- Hoffman, E., McCabe, K., & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review*, *86*, 653-660.
- James, W. (2007/1897). *The will to believe*. New York, NY: Cosimo.
- Horsager, D. (2011). *The trust edge: How top leaders gain faster results, deeper relationships, and a stronger bottom line*. Minneapolis, MN: Summerside Press.
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, *32*, 865-889.
- Johnson, E. J., Shu, S. B., Dellaert, B. G. C., Fox, C., Goldstein, D. G., Häubl, G., ... Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, *23*, 487-504. doi: 10.1007/s11002-012-9186-1
- Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*, 673-676. doi:10.1038/nature03701
- Krueger, J. I. (2013a). Social projection as a source of cooperation. *Current Directions in Psychological Science*, *22*, 289-294. doi: 10.1177/0963721413481352
- Krueger, J. I. (2013b, March). Freedom & power: Feeling powerful has a price. Retrieved from <http://www.psychologytoday.com/blog/one-among-many/201303/freedom-power>
- Krueger, J. I., DiDonato, T. E., & Freestone, D. (2012). Social projection can solve social dilemmas. *Psychological Inquiry*, *23*, 1-27. doi: 10.1080/1047840X.2012.641167
- Luhmann, N. (2000). Familiarity, confidence, trust: Problems and alternatives. In D. Gambietta (Ed.), *Trust: Making and breaking cooperative relations* (pp. 94-107). Oxford, UK: University of Oxford Press.
- Maddux, W. W., Mullen, E., & Galinsky, A. D. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology*, *40*, 461-468.
- McGregor, D. (1960). *The human side of enterprise*. New York, NY: McGraw-Hill.
- Miller, G. (1996). Protean primates: The evolution of adaptive unpredictability in competition and courtship. In A. Whiten & R. W. Byrne (Eds.), *Machiavellian intelligence II: Extensions and evaluations* (pp. 312 - 340). New York, NY: Cambridge University Press.
- Nozick, R. (2001). *Invariances: The structure of the objective world*. Cambridge, MA: Harvard University Press.
- Putterman, L. (2012). *The good, the bad, and the economy*. Minneapolis, MN: Langdon Street Press.
- Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline

- view of trust. *Academy of Management Review*, 23, 393–404.
- Rotter, J. B. (1970). Generalized expectancies for interpersonal trust. *American Psychologist*, 26, 443–452.
- Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. New York, NY: Norton.
- Russell, B. (1938). *Power: A new social analysis*. London, UK: Allen & Unwin.
- Todorov, A. (2008). Evaluating faces on trustworthiness. *Annals of the New York Academy of Sciences*, 1124, 208–224.
- Wedekind, C., & Milinski, M. (1996). Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat. *PNAS*, 93, 2686–2689.

---

Joachim Krueger is a professor of psychology at Brown University. He received his PhD from the University of Oregon and did post-doctoral work at the Max Planck Institute for Human Development and Education in Berlin. His current research interests, besides trust, include self-perception and intergroup relations.



Anthony Evans is an assistant professor of social psychology at Tilburg University. He completed his BA at Northwestern University and received his PhD at Brown University. He currently researches trust in the areas of consumer behavior and social decision making.

