

Current Directions in Psychological Science

<http://cdp.sagepub.com/>

Social Projection as a Source of Cooperation

Joachim I. Krueger

Current Directions in Psychological Science 2013 22: 289

DOI: 10.1177/0963721413481352

The online version of this article can be found at:

<http://cdp.sagepub.com/content/22/4/289>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Association for Psychological Science](http://www.sagepub.com)

Additional services and information for *Current Directions in Psychological Science* can be found at:

Email Alerts: <http://cdp.sagepub.com/cgi/alerts>

Subscriptions: <http://cdp.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Aug 5, 2013

[What is This?](#)

Social Projection as a Source of Cooperation

Joachim I. Krueger

Brown University

Abstract

Social dilemmas, such as the prisoner's dilemma, are often seen as a conflict between rational self-interest and a moral concern for the collective. Contrary to empirical fact, classic game theory says that cooperative behavior cannot occur, and theories of moral sentiment fail to explain how cooperation comes about. Social projection theory offers a fresh look. Seeing decision making as a matter of general inductive reasoning, the theory explains why cooperative behavior occurs, correctly predicts that cooperation will vary with changes in the payoff structure, and shows how a group of self-interested individuals can obtain socially desirable outcomes. Within clearly specified boundary conditions, the theory suggests ways in which social outcomes can be improved beyond current levels.

Keywords

social projection, social dilemmas, interpersonal games, strategic behavior

Whenever I meet people, I feel we already know each other, because we are all the same human beings. Mentally, emotionally, physically we are the same.

~ Tenzin Gyatso, the 14th Dalai Lama,
Providence, Rhode Island,
October 17, 2012

The game: Imagine you and another person have each been given \$5. You may keep the money or give it to the other. Giving is the cooperative choice. If you give the money to the other person, the amount is doubled; that is, if both of you give, you each end up with \$10. Keeping is the defecting choice. If both of you keep, you each get \$5. If one gives and the other keeps, the keeper gets \$15, whereas the giver is left with nothing. This scenario, which is shown in Figure 1, is a “give-some” version (Dawes, 1980) of the familiar *prisoner's dilemma* (Flood, 1952). It is a dilemma because each individual is better off defecting no matter what the other does (Shafir & Tversky, 1992). Yet, if both ignore the rationality of defection, they earn more than if they both defect. Rational self-interest is one horn of the dilemma and the collective interest, or “efficiency,” is the other.

The currency need not be money. Interpersonal respect reflects social power. Respect is worth more to those who receive it than those who give it. Mutual-admiration societies are happiest overall. Yet such

societies are fragile because each individual is motivated to withhold respect in hopes of gaining relative status or out of fear that others will do just that (Krueger, Vohs, & Baumeister, 2008).

Give-some games provide a model for economic and symbolic group behavior. On a national scale, many public goods require a sacrifice from a sufficient number or percentage of individuals. Someone needs to give to charity or public broadcasting, cast votes, pay taxes, and teach the young. Society would be impossible were it not for the many individuals acting irrationally by the lights of game theory (von Neumann & Morgenstern, 1947). If defection dominated the behavior of ordinary people as surely as it dominates the imagination of economists, the curtain on the human tragedy would already have fallen.

Hardin (1968) famously asserted that the gravitational pull of rational defection is so great that the eventual collapse of public resources is as certain as the fall of the tragic hero of classical Greece (see also Diamond, 2005). Hardin was concerned with take-some games, using the plight of open grazing lands as an illustration. In the short term, each herder profits from putting another head

Corresponding Author:

Joachim I. Krueger, Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Box 1821, 190 Thayer St., Providence, RI 02912
E-mail: joachim_krueger@brown.edu

		other person	
		give	keep
you	give	10	0
	keep	15	5

Fig. 1. A symmetrical give-some game. Only the payoffs for Row Player ("you") are shown.

of cattle on the commons, thereby destroying it over the long term. The take-some game is a type of prisoner's dilemma. Its unique feature is that the resource replenishes itself, except that when the number of takers (defectors) surpasses a critical point, everyone suffers. Hardin's words may have been prophetic; overgrazing, overfishing, and overharvesting have reached epic proportions. Likewise, the accelerating pollution of the planet's oceans and atmosphere bespeaks the inexorable logic of en masse defection.

Faced with tragedy, we also observe, however, that generosity, trust, and self-restraint stubbornly survive. This survival presents a challenge. How can we explain why prosocial behavior occurs and where should we look for ways to nurture it? Game theory is no help because it clings to the idea that rational individuals defect (Binmore, 2007), and who wants to purchase cooperation for the price of irrationality? Perhaps there is hope in changing the outcomes such that cooperation becomes the dominating strategy. This was Hardin's (1968) proposal. Following Hobbes (1651/1963), Hardin thought that collective survival requires individuals to endow the state with the power to reward cooperation and to punish defection. Hardin's proposal has practical strength, but it solves the dilemma by defining it away. It does not explain why many individuals cooperate even without being coerced by the Leviathan (i.e., the state).

Why do many people cooperate without being directly rewarded or coerced? These cooperators may hold the key to understanding how cooperation may become more common. Three classes of theory address this challenge. One class accepts the game-theoretic tenet that defection is the rational choice and that cooperation stems from poor thinking or no thinking at all. The second class assumes that cooperators have moral reasons with which they transform and transcend the dilemma's objective payoff structure. The third type of theory assumes that cooperation, when framed as inductive reasoning, can be a rational, self-interested act.

Poor Thinking

When people fail to attend to the consequences of their choices, they might cooperate by mistake. This idea seems extravagant to anyone who has felt the urgency of a social dilemma. Perhaps, then, people recognize the dominance of defection, but their "trembling hands" occasionally hit the COOPERATE button (Cushman, Dreber, Wang, & Costa, 2009). Yet, a theory of random behavior is not much of a theory. A more sophisticated view is that people have learned over time that, compared with defection, cooperation is followed by greater rewards. There is indeed a correlation between behavior (cooperation vs. defection) and outcome unless responses are random. Consider voting. Candidate Oakes beats candidate Roberts if fewer of his supporters abstain (Acevedo & Krueger, 2004). Voting (vs. abstaining) is thus correlated with seeing one's candidate win (vs. lose). Yet, the "poor thinking" theory holds that it is a mistake to base a current decision on past associations (Chater, Vlaev, & Grinberg, 2008). A final lapse of thinking is to mistake a limited game for one that is played forever. Game theorists have proven that in such a game, cooperation can be the smart choice (Fudenberg & Maskin, 1986).¹

Moral Thinking

The theme of the moral approach is that people have values beyond self-interest. They cooperate if they care enough about the social or joint outcome (Van Lange, 1999). The Achilles heel of this idea is that there must be a reason to believe that others will also cooperate. Pure altruists do not care, but they are rare. Many people are conditional altruists; they are willing to give if they think that others will give too. Still, to say that potential cooperators become actual cooperators if they believe that there is a high probability that others will cooperate raises the question of how they came to believe this (Colman, 2003; Gilbert, 1989).

A leaner moral argument is that although some people desire the best joint outcome, they stop short of

estimating the probability that others will cooperate. They decide to do what is necessary for mutual cooperation, knowing that their own choice is not sufficient. This type of reasoning is heuristic (Brandstätter, Gigerenzer, & Hertwig, 2006), but it can be raised to a principle. Someone applying the categorical imperative cooperates to fulfill a duty without regard for what others do (Kant, 1785/1998).

Inductive Thinking

According to the inductive-thinking perspective, cooperation is compatible with rational self-interest. Cognitive errors and moral sentiments may contribute to cooperation, but they are not necessary. My colleagues and I have developed social projection theory to provide a normatively coherent and descriptively adequate framework for cooperation (Krueger, DiDonato, & Freestone, 2012; see also Fischer, 2009, or Pothos & Busemeyer, 2009, for related models). We propose that people use their own choices to predict the choices of others and then select the strategy that is best for them. Projecting their own (intended) choices onto others, they estimate the probability that others will act as they themselves do as being greater than .5. At the limit, the prisoner's dilemma becomes a choice between mutual cooperation and mutual defection. A rational, self-interested player chooses cooperation if it yields a greater personal gain than does defection. The other player's gain need not affect the decision.

Consider the give-some game shown in Figure 1. The expected values are $p \times 10 + (1 - p) \times 0$ and $(1 - p) \times 15 + p \times 5$, respectively, for cooperation and defection, where p is the subjective probability that the other player's choice will be the same as one's own. In this example, a player is indifferent if $p = .75$. Social projection is a rational strategy because there is a majority response for most human traits, preferences, and actions (Humphrey, 1976). Suppose a player believes that 80% of players select a particular option but does not know which option. Considering the payoffs at hand, the player notes that cooperation is more attractive (expected value = $.8 \times 10 + .2 \times 0 = 8$) than defection ($.8 \times 5 + .2 \times 15 = 7$). Why would he or she *not* choose cooperation?

Objections

One argument against projection is that choosing cooperation because it signals a more favorable outcome is magical thinking (Quattrone & Tversky, 1984). When players cannot communicate or influence each other, their choices have no causal force. If they cannot cause others to cooperate, they should defect. The reply is that the association between a player's choices and the

choices of others is a statistical one. As noted by the Dalai Lama (see epigraph), humans share many similarities with one another because of their common ancestry, environment, and culture. Without these shared characteristics, the social or biological sciences would be reduced to case studies.² Recognizing the common causes underlying their own behavior and the behavior of others (Reichenbach, 1951), individuals can cooperate without presuming to cause others to cooperate. These individuals—and the scientists who study them—need only surrender the idea of free will. To paraphrase Schopenhauer (1839/1999), you can choose what you want, but it is not up to you what you want.

A related objection is that only one true probability of cooperation exists. Therefore, a person who makes different probability estimates depending on her own prospective choice must be mistaken. This objection would have a player provide the same estimate irrespective of the strategy being considered for his or her own use. Again, the underlying claim is that individual players can choose their own strategies independently of others and thus independently of that one true probability of cooperation. This claim leads to difficulty.

If a certain proportion of individuals cooperates, who will it be? Not everyone can have independent choice. If, for example, the probability of cooperation is .001, only the first person in a group of 1,000 is free to choose cooperation. Everyone else must defect. Why should this particular player have free choice? If the choices of all others are determined, the first player's choice must also be determined.³ A person who chooses to cooperate could be accused of magical thinking only if this choice were free, but it is not.⁴

Evidence

People cooperate inasmuch as they perceive a high probability that others respond as they themselves do (Acevedo & Krueger, 2005; Fischer, 2009; Krueger & Acevedo, 2007; Yamagishi & Kiyonari, 2000). Moreover, people switch from cooperation to defection if the statistical association between their own choices and the choices of others is broken (Krueger, DiDonato, et al., 2012). As an engine of cooperation, social projection works best if people lack information about others. Consistent with empirical work (Dal Bó, 2005), computer simulations show that as players obtain information about others' choices, projection and cooperation decrease but do not die out (Krueger, Freestone, & DiDonato, 2012).

Studies varying the outcomes also support the projection hypothesis (e.g., Jones, Steele, Gahagan, & Tedeschi, 1968). Consider the give-some game. Many individuals may not project enough (i.e., $p < .75$) to consider cooperation worthwhile. If, however, the transferred money

were tripled, a probability of .66 would be enough. The modified game is “easier” than the original one (Rapoport, 1967), and the projection hypothesis explains why. Theories of cognitive failure, moral sentiment, or social norms do not. Greater projection is good for the group and society. Figure 2 shows the total amounts earned by players in the two give-some games for average projection ranging from 0 to 1. Results were computed from standardized payoffs ($M = 0$, $SD = 1$ within each matrix, easy or hard) to remove spurious effects of overall level of payoff. The graph shows that with increasing projection, groups become more cooperative and thus wealthier, and they do so more strongly for easy than for hard games.

Outlook

Social projection theory makes few assumptions and it overcomes the conventional conflict between personal and social interests. The theory has normative appeal and descriptive fit. It does not prompt the question of the source of the expectation that others will also cooperate. It describes the process as inductive reasoning (Dawes, 1989; Krueger, 2007). Social projection theory is generalizable. Its logic holds in a variety of experimental games and social dilemmas, the prisoner’s dilemma being the most poignant and familiar. Contests among other theories are often considered won or lost by pointing out that a competing theory cannot explain cooperation or coordination in game X.

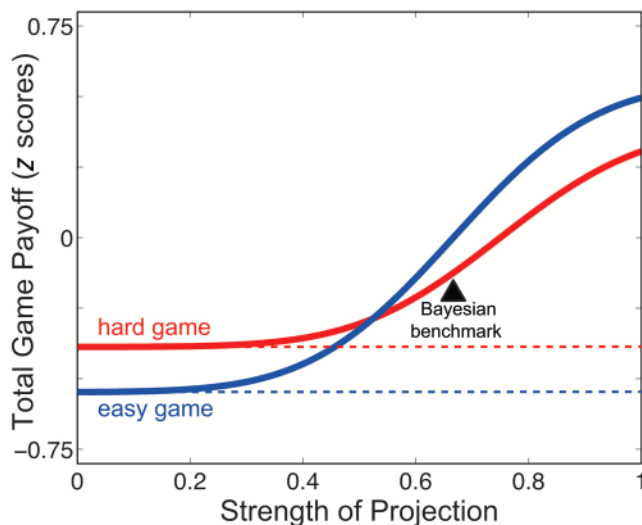


Fig. 2. The summed payoffs as a function of social projection ($SD = .2$ for an average p of .5) in hard and easy give-some games. The Bayesian benchmark refers to the strength of projection ($p = .66$) of rational players who have no information about others. The dashed lines show the summed payoffs garnered by societies of defectors.

Social projection theory is clear about its boundaries. The theory does not apply to games in which players move sequentially or payoffs are asymmetrical. The trust game has both these features (Evans & Krueger, 2011). The second player, or trustee, need not project; the trustee knows whether the first player, or trustor, has invested money in the trustee. The trustor cannot easily project his or her own trust (vs. distrust) onto the trustee’s reciprocation (vs. betrayal) because her situation is different. At most, the trustor can ask what he or she would do in the trustee’s position and project that preference (Krueger, Massey, & DiDonato, 2008). Finally, social projection theory cannot help explain how people find a way to do the opposite of what other players are doing (as desired, for example, in the game of matching pennies).

Social projection works to the degree that individuals lack information about the choices of others. For cooperation to increase, people could be made more ignorant. Such a strategy raises political and ethical concerns. To avoid removal or withholding of information, one can take advantage of social categorization, knowing that social projection is strongest when people see themselves and others as members of the same group (Robbins & Krueger, 2005). Stressing common social categories may be an effective way to make people more cooperative (Singer, 1981). The ability to see the shared humanity in others may be not only a sign of wisdom but also a mark of rationality.

The pessimism of Hobbes and Hardin demands a response. To look for a solution in human stupidity is hardly satisfactory. If cooperation comes from error, are we to adopt a program to make people less rational? Fostering morality and adherence to social norms holds promise (Bicchieri, 2005; Van Vugt, 2009), but these strategies remain linked to notions of threat. People continue to wonder, “If I cooperate out of my goodness, will I be betrayed; if I defect, will I be punished?” In contrast, social projection allows people to form rational expectations of what others will do and act accordingly. The risk of being betrayed is a calculable one.

Recommended Reading

- Fischer, I. (2009). (See References). A series of experiments showing that expectations of self-other similarity positively affect the rate of cooperation.
- Krueger, J. I. (2007). (See References). A broad review of social projection research and its implications for strategic interaction between individuals and between groups.
- Krueger, J. I., DiDonato, T. E., & Freestone, D. (2012). (See References). A target article developing the theme of social projection as a source of cooperation.
- Krueger, J. I., Freestone, D., & DiDonato, T. E. (2012). (See References). A response to nine open peer commentaries to the target article by Krueger, DiDonato, & Freestone.

Declaration of Conflicting Interests

The author declared no conflicts of interest with respect to the authorship or the publication of this article.

Notes

1. But who lives forever to play infinite games?
2. Even game theory would collapse if robbed of its axioms of common knowledge and rationality.
3. To say that this argument against independent choice (i.e., free will) does not apply to infinitely large populations is to raise the question of whether such populations are possible (see Note 2).
4. For a review of arguments against the popular notion of free will, see <http://www.psychologytoday.com/blog/one-among-many/201012/random-walk-through-the-free-will-derness>.

References

- Acevedo, M., & Krueger, J. I. (2004). Two egocentric sources of the decision to vote: The voter's illusion and the belief in personal relevance. *Political Psychology, 25*, 115–134. doi:10.1111/j.1467-9221.2004.00359.x
- Acevedo, M., & Krueger, J. I. (2005). Evidential reasoning in the prisoner's dilemma game. *American Journal of Psychology, 118*, 431–457.
- Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. New York: Cambridge University Press.
- Binmore, K. (2007). *Game theory: A very short introduction*. Oxford, England: Oxford University Press.
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113*, 409–432. doi:10.1037/0033-295X.113.2.409
- Chater, N., Vlaev, I., & Grinberg, M. (2008). A new consequence of Simpson's paradox: Stable cooperation in one-shot prisoner's dilemma from populations of individualistic learners. *Journal of Experimental Psychology: General, 137*, 403–421. doi:10.1037/0096-3445.137.3.403
- Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences, 26*, 139–198.
- Cushman, F., Dreber, A., Wang, Y., & Costa, J. (2009). Accidental outcomes guide punishment in a "trembling hand" game. *PLoS ONE, 4*(8), e6699. doi:10.1371/journal.pone.0006699
- Dal Bó, P. (2005). Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *American Economic Review, 95*, 1591–1604. doi:10.1257/000282805775014434
- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169–193. doi:10.1146/annurev.ps.31.020180.001125
- Dawes, R. M. (1989). Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology, 25*, 1–17. doi:10.1016/0022-1031(89)90036-X
- Diamond, J. (2005). *Collapse: How societies choose to fail or succeed*. New York: Viking.
- 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
- Fischer, I. (2009). Friend or foe: Subjective expected relative similarity as a determinant of cooperation. *Journal of Experimental Psychology: General, 138*, 341–350. doi:10.1037/a0016073
- Flood, M. M. (1952). *Some experimental games*. Research Memorandum RM-789. Santa Monica, CA: RAND Corp.
- Fudenberg, D., & Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica, 54*, 533–544. doi:10.2307/1911307
- Gilbert, M. (1989). Rationality and salience. *Philosophical Studies, 57*, 61–77. doi:10.1007/BF00355662
- Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 243–248.
- Hobbes, T. (1963). *Leviathan* (with an introduction by J. Plamenatz). Cleveland, OH: World Pub. (Original work published 1651)
- Humphrey, N. K. (1976). The social function of intellect. In P. Bateson & R. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge, England: Cambridge University Press.
- Jones, B., Steele, M., Gahagan, J., & Tedeschi, J. (1968). Matrix values and cooperative behavior in the prisoner's dilemma game. *Journal of Personality and Social Psychology, 8*, 148–153. doi:10.1037/h0025299
- Kant, I. (1998). *The groundwork of the metaphysics of morals* (M. Gregor, trans.). New York: Cambridge University Press. (Original work published 1785)
- Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology, 18*, 1–35. doi:10.1080/10463280701284645
- Krueger, J. I., & Acevedo, M. (2007). Perceptions of self and other in the prisoner's dilemma: Outcome bias and evidential reasoning. *American Journal of Psychology, 120*, 593–618.
- 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
- Krueger, J. I., Freestone, D., & DiDonato, T. E. (2012). Twilight of a dilemma: A réplique. *Psychological Inquiry, 23*, 85–100. doi:10.1080/1047840X.2012.666930
- Krueger, J. I., Massey, A. L., & DiDonato, T. E. (2008). A matter of trust: From social preferences to the strategic adherence of social norms. *Negotiation and Conflict Management Research, 1*, 31–52. doi:10.1111/j.1750-4716.2007.00003.x
- Krueger, J. I., Vohs, K. D., & Baumeister, R. F. (2008). Is the allure of self-esteem a mirage after all? *American Psychologist, 63*, 64–65. doi:10.1037/0003-066X.63.1.64
- Pothos, E. M., & Busemeyer, J. R. (2009). A quantum probability explanation for violations of "rational" decision theory. *Proceedings of Biological Sciences, 276*, 2171–2178. doi:10.1098/rspb.2009.0121
- Quattrone, G. A., & Tversky, A. (1984). Causal versus diagnostic contingencies: On self-deception and on the voter's illusion. *Journal of Personality and Social Psychology, 46*, 237–248. doi:10.1037/0022-3514.46.2.237

- Rapoport, A. (1967). A note on the index of cooperation for prisoner's Dilemma. *Journal of Conflict Resolution*, *11*, 101–103.
- Reichenbach, H. (1951). *The rise of scientific philosophy*. Los Angeles: University of California Press.
- Robbins, J. M., & Krueger, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review*, *9*, 32–47. doi:10.1207/s15327957pspr0901_3
- Schopenhauer, A. (1999). *Essay on the freedom of the will*. (Payne E. F. J., trans.). Cambridge, England: Cambridge University Press. (Original work published 1839)
- Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology*, *24*, 449–474. doi:10.1016/0010-0285(92)90015-T
- Singer, P. A. D. (1981). *The expanding circle: Ethics and socio-biology*. Oxford, England: Oxford University Press.
- Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, *77*, 337–349. doi:10.1037/0022-3514.77.2.337
- Van Vugt, M. (2009). Averting the tragedy of the commons: Using social psychological science to protect the environment. *Current Directions in Psychological Science*, *18*, 169–173. doi:10.1111/j.1467-8721.2009.01630.x
- von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly*, *63*, 116–132. doi:10.2307/269588734