PsycCRITIQUES
Contemporary Psychology: APA Review of Books
American Psychological Asso

Trusting Calvin and Hobbes

A review of

Cooperation Without Trust?
by Karen S. Cook, Russell Hardin, and Margaret Levi

New York: Russell Sage Foundation, 2005. 253 pp. ISBN 0-87154-164-5. $32.50

Reviewed by
Joachim I. Krueger

The concept of trust looms large in lay understanding of how interpersonal relationships work, how small groups manage to get things done, and how nations keep from falling apart. The social sciences have dedicated a great deal of effort to the project of rendering a satisfactory definition of trust and of studying how changes in trust affect the well-being of individuals and societies. Trust is often recognized as a critical ingredient of theories of collective action. A common sentiment is that collectives would fare better if their members were only more trusting.[1]

Cook, Hardin, and Levi wrote *Cooperation Without Trust?* as part of a 10-volume series on trust, which they also edit. The length of the series bespeaks the seriousness of the issue, and the involvement of the authors in six of the volumes is a testament to their expertise and dedication to the topic. It is with surprise that one notes the negativity of the message. Trust, they say, has been sloppily defined. When defined crisply, it turns out to make but a minor contribution to the survival of society. Given the Russell Sage Foundation's considerable investment in the matter, this is an anticlimactic message. To evaluate its merit, one may ask whether it goes too far or perhaps not far enough.

Not Gone Far Enough

Cook et al. define trust as "encapsulated interest" that exists when "one party in the relation believes the other party has incentive to act in his or her interest or to take his or her interests to heart" (p. 2). The word *or* is a problem because it allows the trusted party to be self-interested ("has incentive to act") or to be altruistic (takes others' "interests to heart"). A bit later, Cook et al. switch the definition from *or* to *and*: "We trust you because we think you take our interests to heart and encapsulate our interests in your own" (p. 5). Now self-interest and other interest are so blended that an empirical separation is no longer possible. When the trustee acts as hoped, the trustor cannot tell self-interest from other interest. This confusion is a familiar nuisance in the study of altruism. To avoid it, researchers look for behaviors intended to benefit others at a nontrivial cost to the self (Fehr & Fischbacher, 2003).

In the so-called trust game, which is now popular in experimental economics, self-interest and caring for the other can be separated (Berg, Dickhaut, & McCabe, 1995). A trustor receives an endowment and decides what fraction of it to turn over to a trustee. The experimenter multiplies this investment at some preset rate, and the trustee decides how much of it to return to the trustor. In a one-shot encounter, any transfer of funds is irrational according to classic game theory; that is, any return to the trustor squarely violates the trustee's self-interest. Yet exchanges are commonplace, and returns are positively correlated with investments. To break even, though, trustors have to expose themselves to the greatest risk and turn over everything they have (Pillutla, Malhotra, & Murnighan, 2003).

When defining trust as encapsulated interest, one assumes that trustees care not only about their own but also about the trustors' payoffs and, inversely, about the differences between the two. When trustees weight the latter two elements at some fraction of the weight they give to their own payoffs (Van Lange, 1999), trustors should invest all if they suspect that the weight is above one third and otherwise should invest nothing.[2] But how do trustors know what weights the trustees use? There are numerous social-ecological cues, such as kinship, fictive kinship (pp. 97-98), proximity, and familiarity, all of which boil down to perceptions of self-other similarity. The more similar others are in some respect, the more they can be assumed to be similar in other respects, too. Using similarity cues, trustors can simulate in their own minds how much they would return to the other if they were the trustee and then project the result to the other (Krueger & Acevedo, 2005).

Similarity-based predictions are judgments under uncertainty, which Cook et al. want to get rid of. They see encapsulated interest as a property of interpersonal relations where uncertainty is—or at least should be—at a minimum. In one-on-one relationships involving repeated exchanges, trust turns into confident expectations that the other party will cooperate if the first party cooperates. Both parties know that it is in their interest to preserve the relationship, just as they know not to slaughter a lactating cow. However beneficial reciprocal exchange relationships may be, they make it progressively less likely that anyone else's interests need to be taken to heart or that anyone needs to have the experience of trusting another. Ultimately, we can feel comfortable in such relationships because we know with moral certainty that others will do what is best for them and that we will benefit from it. Reflecting this merging of self-interest and other interest and the reduction of uncertainty, the correlation between investments and returns increases when the trust game continues (King-Casas et al., 2005). The assumption of encapsulated interest is no longer necessary; it is parsimonious to refer to overlapping interests.

Cook et al.'s definition of cooperation hinges on its distinction from mere coordination, where the latter is what one seeks in, for instance, assurance games. Once encapsulated interests are set aside, the distinction between cooperation and coordination also loses force. Participants in mutually beneficial exchange relationships have found a way to maximize their returns over the long run by following their self-interests. The interests of the other are met in passing, and it is sufficient to talk of mutual coordination. If overlapping interests and successful coordination can explain the survival of mutually beneficial exchange relationships, the attempt to downplay the role of trust and cooperation does not go far enough. What is called trust is perhaps better understood as just another kind of inductive inference, like an expectation, a prejudgment, or a bet.

Gone Too Far

The distinctions between encapsulated and overlapping interests and between cooperation and coordination seem like definitional quibbles in light of the book's major message, which is that relational trust cannot matter at the level of society. The given reason is that "trustworthiness and reliability are generally too important in society to leave to the vagaries of interpersonal relations or to the unpredictability of individual behavior" (p. 31). Cook et al. present no evidence to support this claim. Instead, they tie the claim to the definition of trust as encapsulated interest. Then, it follows, of course, that "trust cannot play a large role because the relationships between all the relevant parties cannot be rich enough to ground trust" (p. 190).

As followers of Hobbes, Cook et al. decry the craziness of trusting big government, petty bureaucrats, or anyone in between. How, then, is society possible if distrust prevails outside the confines of tightly knit relationships? Cook et al. consider a myriad of informal and semiformal ways cooperation can be induced and maintained. They emphasize the role of "informal contractual dealings," "reputational effects," "background protections," and "communal norms," which are best realized in midlevel organizations and networks that provide "webs of group affiliations" (p. 188) and "networks of control" (p. 190).

Mechanisms of external regulation are important for social cohesion and productivity. They are, however, poorly differentiated from one another, and they induce not only cooperation but all kinds of yielding behavior, from conformity to compliance to obedience. If we agree that a free society cannot function without the consent of its citizens, why struggle so hard to marginalize trust? The struggle is a lost cause, anyway. Trust keeps creeping back in at every turn. Chapter 7, for example, raises the question of how organizations can induce cooperation or other yielding behaviors among their employees. Surely, "trust is not enough" (p. 146), but who wants the coercion of a forced-labor camp? The use of soft surveillance techniques is partially successful but limited by the power differential (Chapter 3).[3] Why should employees trust organizations that reserve the right to punish them? Direct demands of the "Trust us!" variety certainly do not help either. All they do is signal that management does not trust the employees to trust it. As in the experimental trust game, an initial voluntary investment is necessary to jump start a mutually profitable relationship. Chapter 9 shows how exchange relationships emerge when other modes of social control fail. During revolutions and mass migrations, people are most likely to grow exchanges from the roots of kinship or familiarity.

Trying to reduce the role of trust, Cook et al. might have gone too far. In the early stages of interpersonal relations, trust is necessary to set mutually beneficial exchanges in motion and, compared with social arrangements for the external regulation of behavior, it is less vulnerable to sudden upsets. When such upsets happen, many of these regulatory arrangements, such as communal norms and carefully crafted reputations, can only be rebuilt through networks of trust-based relationships.

Conclusion

The suggestion that trust is a special kind of inductive reasoning has a Calvinist flavor. Trust involves a level of certainty that lies between hope and faith. Calvinists engage in good works and vigorous business activity not because they believe these efforts will be rewarded but because they believe the efforts are diagnostic of the desired state of grace, which, in turn, is revealed in earthly success. Likewise, trustors in experimental games cannot make the trustees reward them for their confidence in them. Individual voters in national elections cannot make millions of like-minded others stream to the polls, but they can trust that their own efforts are diagnostic of the popular will (Acevedo & Krueger, 2004). Trust in the similarity of others to the self, and thus in their willingness to reciprocate, creeps back in. Alas, these inductions cannot be perfect. They can reduce but not solve the dilemma of collective action. As long as free riding is an option, Hobbes's Leviathan cannot be chained. The choice between Calvin and

Hobbes is a false one. We need both to survive.

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. PsycINFO Article

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

Fehr, E., & Fischbacher, U. (2003, October 23). The nature of human altruism. *Nature*, *425*, 785-791. PsycINFO Article

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005, April 1). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, *308*, 78-83. PsycINFO Article

Krueger, J. I., & Acevedo, M. (2005). Social projection and the psychology of choice. In M. D. Alicke, D. Dunning, & J. I. Krueger (Eds.), *The self in social perception* (pp. 17-41). New York: Psychology Press.

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. PsycINFO Article

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. PsycINFO Article

[1]The understanding that cooperation is moral but stupid, whereas defection is selfish and smart, reflects certain habits of thought rather than intrinsic properties of behavior. When a riot is afoot, for example, those who selflessly cast the first stone do not act morally.

Being selfish, smart, and evidently moral, the free riders stay home, hoping that others will put themselves at risk for the benefit of the community. When defection is good and cooperation is bad, one hopes that when they give a war, no one shows up. It is even possible for the same behavior to be cooperative or defecting. Producing many offspring, for example, can be seen as a selfish choice that threatens the viability of the whole community. When two competing populations try to outbreed each other, however, members of each community can expect their fellow members to cooperate toward the group goal by being prolific.

[2]Formally, the trustee's effective payoff is $mpE - qmpE + w(qmpE + E - pE) - w|mpE - qmpE - qmpE - E + pE|$, where E is the trustor's endowment, pE is the proportion invested, m is the multiplier applied to the investment, $qmpE$ is the proportion of the multiplied investment returned to the trustor, and w is the weight representing how much the trustee cares about the trustor's objective payoff and about the difference between the trustor's and the trustee's objective payoffs.

[3]If organizations can worry about being trusted by their employees, perhaps governments should learn to worry about being trusted by their citizens.

PsycCRITIQUES February 22, 2006 Vol. 51 (8), Article 17

1554-0138