

Psychological Science

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

Why So Cynical? : Asymmetric Feedback Underlies Misguided Skepticism Regarding the Trustworthiness of Others

Detlef Fetchenhauer and David Dunning

Psychological Science published online 8 January 2010

DOI: 10.1177/0956797609358586

The online version of this article can be found at:

<http://pss.sagepub.com/content/early/2010/01/08/0956797609358586>

Published by:



<http://www.sagepublications.com>

On behalf of:



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

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

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

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

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

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

Why So Cynical? Asymmetric Feedback Underlies Misguided Skepticism Regarding the Trustworthiness of Others

Detlef Fetchenhauer^{1,2} and David Dunning³

¹University of Cologne, ²University of Groningen, and ³Cornell University

Psychological Science

XX(X) 1–5

© The Author(s) 2010

Reprints and permission: <http://www.sagepub.com/journalsPermissions.nav>

DOI: 10.1177/0956797609358586

<http://pss.sagepub.com>

 SAGE

Abstract

People tend to grossly underestimate the trustworthiness of other people. We tested whether this cynicism grows out of an asymmetry in the feedback people receive when they decide to trust others. When people trust others, they painfully learn when other people prove to be untrustworthy; however, when people refrain from trusting others, they fail to learn of instances when the other person would have honored their trust. Participants saw short videos of other people and had to decide whether to trust each person in an economic game. Participants overall underestimated the trustworthiness of the people they viewed, regardless of whether they were given financial incentives to provide accurate estimates. However, people who received symmetric feedback about the trustworthiness of others (i.e., who received feedback regardless of their own decision to trust) exhibited reduced cynicism relative to those who received no feedback or asymmetric feedback (i.e., who received feedback only after they trusted the other person).

Keywords

trust, economic games, social prediction, cynicism, experience sampling

Received 4/11/09; Revision accepted 6/30/09

People can be cynical to a fault. They underestimate how often others respond generously to requests for help (Flynn & Lake, 2008) and overestimate how much others' attitudes and actions are driven by selfish concerns (Miller, 1999). To be sure, there is contrary evidence showing that people can be roughly realistic in anticipating the altruism of others (Balcetis & Dunning, 2008; Balcetis, Dunning, & Miller, 2008; Epley & Dunning, 2000, 2006), but an increasing body of evidence suggests that when people are contemplating whether they should rely on the kindness of strangers, they suspect those strangers will prove more selfish than actually is the case.

We have previously shown this cynicism most clearly in experiments using the economic paradigm known as the “trust” or “investment” game (Berg, Dickhaut, & McCabe, 1995; Fetchenhauer & Dunning, 2009). In the game, the truster is given money that can be kept or handed to a completely random and anonymous stranger, the trustee. If the truster hands his or her money over, the amount of money is quadrupled (e.g., \$5 becomes \$20), and trustees have two options: They can either split the money evenly between themselves and the truster (e.g., give \$10 back and keep \$10 for themselves), or they can keep all the money for themselves.

In a number of studies, we have shown that the vast majority of trustees honor the trust that is offered them, giving money back even when their identities are anonymous and they are under no compulsion to act generously (Fetchenhauer & Dunning, 2009). However, most trusters severely underestimate their fellow participants' trustworthiness. Although 80% to 90% of trustees honor trust, trusters on average estimate that this rate will be only 45% to 60%.

Such cynicism may matter, in that it leads people to refrain from trusting, and thus pass up likely monetary gains (Fetchenhauer & Dunning, 2009). For example, in versions of the game in which people can decide the exact amount of money to transfer to the trustee, many people pass along only a little money. This causes trustees to pass back little, although they would have been quite generous if they had been trusted more completely (Pilutla, Malhotra, & Murnighan, 2003). Similarly, when trusters impose possible penalties if trustees are not

Corresponding Author:

Detlef Fetchenhauer, Department of Economic and Social Psychology, University of Cologne, Herbert-Lewin Str. 2, 50931 Cologne, Germany
 E-mail: detlef.fetchenhauer@uni-koeln.de

generous enough in return, trustees respond by passing back significantly less than they would otherwise (Fehr & Rockenbach, 2002; Malhotra & Murnighan, 2002).

The research reported here investigated how such skepticism toward others can be explained. We focused on two possible determinants. First, participants may not be sufficiently motivated to provide accurate estimates of trustworthiness. Behavioral economists would argue that participants provide valid estimates only if they have a sufficient material incentive to do so, although this argument cannot explain why estimates are so biased in one direction.

Second, when people decide whether to trust, they can make two mistakes. They can trust someone whose intentions are actually harmful, or they can refuse to trust a person who would actually reciprocate that trust (see Table 1). The chance that life informs them about these two mistakes is asymmetric. When people trust another person and that person betrays their trust, people become painfully aware of that betrayal. However, if people distrust another person, they never give that person a chance to act in a trustworthy way. Thus, people preclude themselves from learning when others, despite expectations, might prove trustworthy.

In essence, we argue for an *experience-sampling* explanation for unwarranted cynicism in situations involving trust. When people trust and that trust is exploited, they become more cynical. However, when they mistakenly fail to trust a trustworthy person, they avoid the experience that would provide feedback to correct that mistake. Thus, their overall impression of human nature is left overly cynical. Such experience-sampling explanations have been offered for a variety of phenomena in social interaction. Researchers have proposed that experience sampling may explain negativity biases in opinions about attitude objects (Fazio, Eiser, & Shook, 2004) and impressions of other individuals and social groups (for a review, see Denrell, 2005). Experience-sampling explanations have also been proposed to explain how people learn that cooperation in a prisoner's dilemma game is rewarding, even though economic analysis indicates that cooperation should be avoided (Chater, Vlaev, & Grinberg, 2008).

But the empirical base for experience-sampling theorizing is thin (Denrell, 2005; Smith & Collins, 2009). We undertook this study to explore whether such an account could explain the overall cynicism we see in the trust game. We examined

whether experience sampling could influence the beliefs people hold about human nature, and not just about individuals, as has been the focus of past studies. Additionally, we explored whether experience sampling influences choices in social interactions that carry real outcomes, and, thus, whether it ultimately alters the outcomes that participants obtain. We focused on the trust game because of its increasing and broad prominence in research and theory across economics, sociology, and psychology (for a review, see Dunning & Fetchenhauer, in press).

Participants viewed videos of 56 individuals and indicated whether or not they would trust each of them. We varied whether there was a financial incentive for making accurate estimates and the amount of feedback they received after they had made their choices. Some participants received no feedback about the accuracy of their judgments; others received feedback only when they decided to trust the other individual, the situation that we presume occurs in real life; and still others received feedback for all interactions regardless of their decision, so that the informational asymmetry we propose exists in real life was removed. If asymmetry is responsible for the cynicism seen in past studies, eliminating it should eradicate the cynicism.

Method

Stimuli

Fifty-six individuals were videotaped at the University of Groningen. Each individual sat in front of a white wall and introduced himself or herself. Each video lasted 10 s, with the sound switched off the entire time. After they were videotaped, the individuals were asked to make decisions in a number of experimental games (see Fetchenhauer, Pradel & Groothuis, in press). One of these decisions was what they would do as a trustee in the trust game.

Procedure

We described the trust game, similar to the one described in the introduction, to 120 University of Cologne students. They were told that they would see 56 short, prerecorded videos, each presenting a person who had played the role of the trustee in the game. Participants themselves were to play the role of the truster. They would be given €7.50 and asked whether they wanted to keep the money or give it to the person on the videotape. If they gave it to the person on the videotape, the amount would be inflated to €30. Participants were told that the person on the videotape had already indicated whether he or she would split the money with the truster (giving €15 back) or keep it all. Before watching any of the videos, participants estimated the percentage of the 56 trustees they thought would reciprocate trust (i.e., hand money back).

Participants then viewed the short videos. After viewing each one, participants predicted whether they thought that person

Table 1. Possible Mistakes in Decisions to Trust or Not Trust

Is interaction partner trusted?	Is interaction partner actually trustworthy?	
	Yes	No
Yes	Warranted trust	Unwarranted trust
No	<i>Unwarranted distrust</i>	<i>Warranted distrust</i>

Note: Cells in boldface reflect situations in which individuals receive feedback for their decisions in real-world contexts. Cells in italics reflect situations in which no feedback is received in real-world contexts.

would prove trustworthy (“Do you think that this stimulus-person split the money or did that stimulus-person keep all of it?”). They also decided whether they would give that person their initial €7.50 (“Do you hand over your €7.50 to that stimulus-person?”). Participants were told that one of their decisions would be played for real.

After viewing all the videos, participants again estimated the overall percentage of trustees who had given money back versus kept it.

Incentives and feedback

In a 2×3 between-subjects design, we varied financial incentives and feedback. In the financial-incentive condition, participants were told they would receive €0.50 for each trustee whose behavior they guessed correctly. They would forfeit €0.50 for each wrong guess. If they made more wrong than right guesses, they would not have to pay any money to the experimenter. No incentive was mentioned in the no-financial-incentive condition.

Feedback was manipulated across three conditions. In the noncontingent-feedback condition, participants were informed about the actual decision of each trustee immediately after they decided whether or not to trust that person (i.e., participants received feedback regardless of their own decision). In the contingent-feedback condition, participants received such feedback only when they decided to trust the trustee. In the no-feedback condition, participants received no feedback.

Results

Consistent with past work, our results showed that participants displayed undue cynicism. Before viewing any videotapes, participants, on average, stated that 52.1% ($SD = 17.8$) of the trustees would prove trustworthy. This percentage was well below the actual value of 80.4%, $t(119) = -17.14, p < .001, p_{rep} > .99$.

Playing the trust game allowed participants to rid themselves of this cynicism—but whether they did depended on the type of feedback they received, as revealed in a number of 2 (incentive) \times 3 (feedback) analyses of variance. Essentially, contingent feedback had no impact on participants’ perceptions or behavior, but noncontingent feedback did. It prompted participants’ estimates to be more optimistic, their behavior to be more trusting, and their wallets to be more full.

Regarding predictions about individual trustees, participants in the contingent-feedback condition estimated that 60.4% of trustees were trustworthy. This value was close to the 56.9% estimated by participants in the no-feedback condition. However, in the noncontingent-feedback condition, participants on average estimated that 71.0% of trustees were trustworthy, and predictions of trustee trustworthiness varied significantly by feedback condition, $F(2, 114) = 15.31, p < .01, p_{rep} > .99$. This led to increased acts of trust, as reflected by participants’ actual decisions. In the noncontingent-feedback condition, participants handed their money to 70.1% of all

trustees, whereas in the contingent- and no-feedback conditions, participants handed their money to 58.5% and 56.8% of trustees, respectively, $F(2, 114) = 8.11, p < .01, p_{rep} > .99$. This greater level of trust had an impact on financial outcomes. Participants in the no-feedback condition earned €6.36 ($SD = €1.29$) on average, about the same as participants in the contingent-feedback condition ($M = €6.53, SD = €1.35$). Participants in the noncontingent-feedback condition earned substantially more ($M = €7.67, SD = €1.11$), $F(2, 113) = 12.98, p < .001, p_{rep} > .99$.

Further analyses revealed that participants in the noncontingent-feedback condition learned quickly about others’ trustworthiness, and adapted their behavior accordingly. We split the 56 decisions into four sequential blocks of 14, and found that both predictions of trustee behavior and decisions to trust changed significantly only between the first and second blocks, $t_s(119) = 3.69$ and 3.66 for predictions and decisions, respectively, both $p_s < .01, p_{rep_s} > .99$. After the second block, predictions of trustworthiness and decisions to trust continued at a stable rate that was higher than the rate during the first block.

Finally, it appears that participants did gain an accurate impression of the trustworthiness of the 56 trustees after receiving noncontingent feedback. Our last analysis focused on participants’ estimates of the overall percentage of trustees they believed were trustworthy. The average estimate was 79.1% in the noncontingent-feedback condition, 62.6% in the contingent-feedback condition, and 57.6% in the no-feedback condition, $F(2, 114) = 28.99, p < .001, p_{rep} > .99$. Thus, the difference between the estimate and the actual value of 80.4% was not significant in the noncontingent-feedback condition, $t = -0.8$, but was significant when participants received only contingent feedback, $t(114) = -6.243, p < .001, p_{rep} > .99$, or no feedback, $t(114) = -8.97, p < .001, p_{rep} > .99$.

Post hoc tests (least significant difference) indicated that for all four dependent variables (predictions of individual trustee behavior, decisions to trust, and estimates of trustworthiness before and after watching the videos), the only significant differences were between the noncontingent-feedback condition and the two other conditions ($p_s < .05$). Financial incentives had no impact on accuracy, except for a marginal impact on decisions to trust, $F(1, 119) = 3.54, p = .06, p_{rep} = .91$. Participants handed their money to trustees slightly more often in the financial-incentive condition ($M = 64.6\%, SD = 17.0$) than in the no-financial-incentive condition ($M = 58.0\%, SD = 17.1$).

Discussion

We found that people tend to substantially underestimate the trustworthiness of others, a result echoing past work (Fetchenhauer & Dunning, 2009). We tested two explanations for this unwarranted cynicism. One explanation was that people do not provide accurate assessments unless material incentives motivate them to do so. However, we found little evidence that providing a motivation for accuracy rids people of their cynicism.

The degree to which participants underestimated trustees' trustworthiness was only marginally influenced by whether participants had a financial incentive to produce a valid estimate, and this influence was found for only one measure.

Our experience-sampling hypothesis was based on the notion that experience informs people in an asymmetric way about the different kinds of mistakes they make about the trustworthiness of others. Our results were in line with this reasoning. The degree of skepticism our participants displayed did not differ between the no-feedback condition and the contingent-feedback condition, in which participants received feedback only when they decided to trust the trustee. We argue that this feedback condition comes closest to mirroring people's experience in real life, and it seems as though people fail to learn about others' true level of trustworthiness because of this asymmetric experience.

However, when we gave participants feedback about the trustees' trustworthiness irrespective of participants' decisions, they adjusted their estimates and decisions toward the actual level of their interaction partners' trustworthiness. Perhaps a lack of symmetric feedback explains why people generally see others as driven more by self-interest than they really are (Miller, 1999). People are exposed to instances of surprising self-interest (e.g., when they trust and that trust is violated), but not to instances of surprising selflessness. Of course, we hasten to add that this experience-sampling explanation is only one of many possible explanations for undue cynicism. Other psychological processes may also be in play.

Our findings leave us with two possible lessons. First, over the past 35 years, psychological theorizing has sought to explain judgmental bias by emphasizing flaws in human cognition (e.g., Gilovich, Griffin, & Kahneman, 2002; Nisbett & Ross, 1980). We suggest, as have others, that biases in social belief may arise not from flawed cognition, but rather from flaws or gaps in the information people acquire (see Denrell, 2005; Dunning, 2005; Smith & Collins, 2009). Thus, the flaw is not in how people think about information; it is in the information they think about.

Second, our findings imply a self-fulfilling prophecy in dealing with one's social environment. In high-trust environments (e.g., an intimate relationship or a work organization), people will tend to trust each other and be rewarded (at least if such trust is generally justified). In low-trust environments, people will not trust each other and will never learn if their distrustfulness is actually warranted. Thus, people may learn that their friends and intimates are terrific people—and fail to learn just how terrific everyone else is.

Acknowledgments

We thank Tobias Große and Christian Krebeck for assistance in collecting the data.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interests with respect to their authorship and/or the publication of this article.

Funding

This research was supported financially by the University of Groningen and by National Science Foundation Grant 0745806 (awarded to D. Dunning).

References

- Balceris, E., & Dunning, D. (2008). A mile in moccasins: How situational experience reduces dispositionism in social judgment. *Personality and Social Psychology Bulletin, 34*, 102–114.
- Balceris, E., Dunning, D., & Miller, R.L. (2008). Do collectivists “know themselves” better than individualists? Cross-cultural investigations of the “holier than thou” phenomenon. *Journal of Personality and Social Psychology, 95*, 1252–1267.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*, 122–142.
- 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.
- Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review, 112*, 951–978.
- Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychological Press.
- Dunning, D., & Fetchenhauer, D. (in press). Why people trust. In D. Dunning (Ed.), *Social motivation*. New York: Psychological Press.
- Epley, N., & Dunning, D. (2000). Feeling “holier than thou”: Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology, 79*, 861–875.
- Epley, N., & Dunning, D. (2006). The mixed blessings of self-knowledge in behavioral prediction: Enhanced discrimination but exacerbated bias. *Personality and Social Psychology Bulletin, 32*, 641–655.
- Fazio, R.H., Eiser, J.R., & Shook, N.J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology, 87*, 293–311.
- Fehr, E., & Rockenbach, B. (2002). Detrimental effects of sanctions on human altruism. *Nature, 422*, 137–140.
- Fetchenhauer, D., & Dunning, D. (2009). Do people trust too much or too little? *Journal of Economic Psychology, 30*, 263–276.
- Fetchenhauer, D., Pradel, J., & Groothuis, T. (in press). Not only states but traits: Humans can identify permanent altruistic dispositions in 20 seconds. *Evolution & Human Behavior*.
- Flynn, F.J., & Lake, V.K.B. (2008). If you need help, just ask: Underestimating compliance with direct requests for help. *Journal of Personality and Social Psychology, 95*, 128–143.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.
- Malhotra, D., & Murnighan, J.K. (2002). The effects of contracts on interpersonal trust. *Administrative Science Quarterly, 47*, 534–559.

- Miller, D.T. (1999). The norm of self-interest. *American Psychologist*, *54*, 1053–1060.
- Nisbett, R.E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Pilutla, M.M., Malhotra, D., & Murnighan, J.K. (2003). Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology*, *39*, 448–455.
- Smith, E.R., & Collins, E.C. (2009). Contextualizing person perception: Distributed social cognition. *Psychological Review*, *116*, 343–364.