

Retrospection in Social Dilemmas: How Thinking About the Past Affects Future Cooperation

Craig D. Parks
Washington State University

Lawrence J. Sanna
University of North Carolina at Chapel Hill

Donelle C. Posey
Washington State University

This article reports 2 studies investigating the effects of retrospective thought on future cooperation in social dilemmas. Some general theories of cooperation presume, but have not tested, whether retrospection has impact: People may think about the choices they could have made instead, realize that cooperation would have produced larger outcomes, and change their strategy as a result. Across both studies, the authors show that rate of future cooperation is directly related to the number of best-case scenarios and inversely related to the number of worst-case scenarios generated. The 2nd study also shows that the number and type of retrospective thoughts generated can be predicted from the person's social value orientation.

A social dilemma is a situation of interdependence in which there is a conflict between individual and collective gain: The action that has the potential to produce the best personal outcome is not the one that will produce the best collective outcome. Further, if all persons act in their own best interests, then everyone will be worse off than if all had acted in the collective best interest. The classic example of a social dilemma is the prisoner's dilemma, and more recently researchers have studied resource dilemmas and public goods situations, both of which retain most of the features of a social dilemma and have the added attraction of well modeling a number of real-world events (for a summary of this work, see Komorita & Parks, 1995, 1996).

Researchers in this area are typically interested in understanding the dynamics of cooperative (i.e., group-regarding) choice. In particular, theorists are interested in the question of how to encourage greater rates of cooperation in people. Many potential interventions have been investigated, but the field is generally lacking in theoretical models that can pull together the results from these many studies and explain how and why people decide to become more or less cooperative (Komorita & Parks, 1995). There are two general models that have received some attention: Axelrod's (1984) model of effective choice, and Camerer and Ho's (1998) experience-weighted attraction learning model. Although the two models differ in many ways, they have in common the

assumption that participants in a social dilemma will periodically compare their outcomes to what they might have earned had they been using alternate choice strategies. Through this comparison, they will discover that they would have accumulated more if they had been more cooperative (presuming, of course, that the person is not already perfectly cooperative), and as a result they will become more cooperative in the future.

This article reports on two studies that test this assumed impact of mental simulation of alternative courses of action. As shown below, research on counterfactual thinking suggests that people who do engage in such retrospection will not necessarily focus on how things could have been better and that people who instead focus on how things could have been worse will not alter future rates of cooperation for the better. Thus, our goal was to understand exactly how mental simulation relates to future cooperation.

Two Models of Cooperation

Axelrod's (1984) model of effective choice grew out of his computer simulations of choice strategies in two-person prisoner's dilemmas. Basically, the simulations demonstrated the broad effectiveness and consistency of the tit-for-tat strategy at accumulating outcomes across a wide variety of opponent strategies. Axelrod used these results to also formulate a set of recommendations for how to choose effectively in a mixed-motive situation. In brief, he recommended that one (a) be unconcerned with how well others are doing, (b) never be the first to make an uncooperative choice, (c) reciprocate both cooperative and uncooperative actions, and (d) avoid overly clever or intricate choice strategies. It is this first recommendation that is of interest here. Axelrod argued that an examination of others' outcomes relative to one's own provokes feelings of envy, which in turn lead to vengeful, uncooperative behavior. (To what extent such comparisons actually occur in and are destructive to a social dilemma have been docu-

Craig D. Parks and Donelle C. Posey, Department of Psychology, Washington State University; Lawrence J. Sanna, Department of Psychology, University of North Carolina at Chapel Hill.

We thank Caroline Roggenbuck, Asako Sasaki, and Jamie Severson for their assistance in data collection.

Correspondence concerning this article should be addressed to Craig D. Parks, Department of Psychology, Washington State University, P.O. Box 644820, Pullman, Washington 99164. E-mail: parkscd@mail.wsu.edu

mented elsewhere; Parks, Rumble, & Posey, 2002.) Axelrod said that people should instead compare their outcomes against those that they would have received had they used other choice strategies. To quote Axelrod directly, “A better standard of comparison is how well you are doing relative to how well someone else could be doing in your shoes. Given the strategy of the other player, are you doing as well as possible?” (p. 111). The presumption is that, on such reflection, the person will realize the value of long-term cooperation and become more cooperative in the future.

The experience-weighted attraction (EWA) learning model (Anderson & Camerer, 2000; Camerer & Ho, 1998, 1999a, 1999b) argues that the probability with which a person will make a particular choice in a mixed-motive situation is a function of the attraction, $A_i^j(t)$, of that choice, in terms of its ability to return a desirable outcome,¹ immediately after the current interaction concludes. Attraction is affected by the number of recent past experiences, $N(t)$, to which one is still responding (which are not necessarily equivalent to the simple number of past trials), as well as two continuously modified parameters: An attraction depreciation parameter, ϕ , which reflects the degree to which original predispositions are decaying, and an imagination parameter, δ , which reflects the person’s ability to simulate outcomes that could have resulted had a different choice been made. $N(t)$ is itself modified by an experience depreciation parameter, ρ , which can be thought of as a decline in the value of a specific past event (e.g., how someone behaved 10 trials ago) for predicting future events.² When $t = 0$, N and A_i^j represent pregame experience, with N reflecting the expected reinforcement value of an outcome, and A_i^j reflecting initial expectations about how others will behave. The size of $N(0)$ indicates how strongly the reinforcement value will persist in the face of actual experience. A large value means that the person will continue to be attracted to a specific outcome, even if the actions of others are making it difficult to achieve the outcome, and a small value means that the outcome will quickly lose its attractiveness if it proves hard to earn.

This model has much to offer social psychologists, but for our purposes here we focus solely on the imagination parameter. It acts as a weighting factor for the payoff imagined to be associated with the foregone choice and affects the probability of repetition of a choice, in that if the imagined outcome is large and the ability to imagine it is high (i.e., δ is large), all else being equal, the likelihood of repeating the previous choice will decrease. Thus, as with Axelrod’s (1984) model, the presumption here is that an uncooperative person’s mental simulation of better payoffs associated with mutual cooperation will motivate him or her to adopt a cooperative strategy in the future. The influence of mental imagery in the EWA model is more complex—the person has to be willing to use information about unchosen alternatives and attend to the actions of others (i.e., ρ is large)—but mental simulation is a key element in the model. Further, the implication of both Axelrod’s and the EWA models is clear: Thinking about alternative courses of action can make one a more cooperative person.

Mental Simulation

The focus on imagining alternative courses of action in the Axelrod (1984) and EWA models suggests that another area of research, on mental simulations more generally, may greatly inform an understanding of cooperative and competitive behavior

(for reviews, see Sanna, 2000; Sanna, Stocker, & Clarke, 2003). In particular, research on counterfactual thinking—thoughts about “what might have been”—has demonstrated that mentally simulated alternatives occur in one of two directions. *Upward counterfactuals* are mentally simulated past outcomes that are better than what actually happened, whereas *downward counterfactuals* are mentally simulated past outcomes that are worse than what actually happened. For example, a person might think upward counterfactually, “If only I would have performed Action X, I might have done better.” By contrast, a person might also think downward counterfactually, “At least I got something—it could have been worse.” This distinction is important because both Axelrod and EWA focus only on upward mental simulations. People will ask how they could have earned better outcomes and will discover that such a gain is most likely to occur via mutual cooperation.

When people imagine alternative courses of action, however, there is no guarantee that they will focus on how things could have been better. People who retrospect may focus on poorer rather than better alternate states of nature, and this may allow them to feel better at least in the short term (Sanna, Turley-Ames, & Meier, 1999), although this also suggests that those who focus on worst-case scenarios will not necessarily become more cooperative in the future. That is, downward counterfactuals may encourage maintenance of old behaviors rather than adoption of new ones. If this is true, it follows that although mental simulations of better scenarios may indeed promote more frequent cooperation, as the Axelrod and EWA models suggest, simulations of worse scenarios may actually inhibit the development of future cooperation.

How might counterfactual thoughts provoke more or less cooperation? One possibility is that the type of thought instills a particular mood, which in turn affects rate of cooperation. Markman, Gavanski, Sherman, and McMullen (1993) showed that generation of upward counterfactuals can lead to negative moods (oriented around dissatisfaction), and generation of downward counterfactuals can lead to positive moods. Mood influences on cooperative choice have so far proven difficult to untangle (Hertel, 1999). However, research on mood and motivation has shown that, when people are oriented toward performance goals, negative moods are taken as indicative of a failure to achieve the goal and enhance performance, whereas positive moods are interpreted as meaning that the goal has been achieved and produce performance decrements (Martin, Ward, Achee, & Wyer, 1993). Dilemma settings typically emphasize “doing well,” so it is entirely plausible that participants are performance oriented. It could thus be that thinking of how things could have been better in a social dilemma produces a negative mood, which is taken to mean that the goal has not been achieved, which spurs more cooperation, whereas thinking of how things might have been worse produces a positive mood, which leads to the inference that one’s performance has

¹ The EWA learning model assumes that people are not necessarily driven to always realize the best outcome. Rather, people are assumed to want satisfice and to strive for any outcome that is at least somewhat attractive.

² When $\delta = 0$ (i.e., the person is unconcerned with unobserved outcomes) and $\rho = 0$ (i.e., the person is not influenced by the past actions of others), the model is equivalent to a pure reinforcement model.

been satisfactory, which leads to no change or a decline in rate of cooperation.

The basic purpose of the two studies reported here was to establish that type of counterfactual thought is related to rate of cooperation. Demonstrating such a connection creates a rationale for testing our hypothesized mechanism in future studies. Our first study was designed to examine the types of thoughts people have when directed to think about a recently completed mixed-motive interaction and to see if these thoughts could predict future rates of cooperation. Our prediction was that the number of upward counterfactuals would be positively related and the number of downward counterfactuals negatively related to the rate of future cooperative choice. The purpose of our second study was to test whether a person's social value orientation would predict the direction of mentally simulated alternatives and subsequently lead to differential rates of cooperation.

Study 1

Participants

The participants were 80 students enrolled in an introductory psychology course. Participation was in partial fulfillment of a course requirement.

Method

Students were run in sessions of 4 persons. On entering the lab, they were seated in high-walled cubicles that contained the matrix for a two-person prisoner's dilemma game (shown in Figure 1) and a tripole switch-box marked "A" at the top, "OFF" in the middle, and "B" at the bottom. They were told that they would be playing many trials (30, though this number was left unspecified) of a game with another, unidentified student. On each trial, they were to choose between "A" and "B." At the same time, the other person would also be making the same choice. The combination of the two choices would determine how many points each person received on that trial. Examples were provided of how to read the matrix. Students were told to make their choices by moving the switch to either the "A" or "B" position. The switches were connected to a central light box, seen only by the experimenter. Moving the switch to a particular position illuminated a specific light that indicated whether the student had chosen "A" or "B." (The "OFF" position turned off all lights.) In this way, each person could make a choice without any other student being aware of his or her choice. After everyone had registered a choice, the experimenter gave each person a slip of paper with his or her choice and point total printed on it. Students were told that at the end of the semester, three people's names would be chosen at random from the master list, and each person's total points would be converted to cash at the rate of 10 cents per point. As a result, although each person had an incentive to accumulate as many points as possible, because winners were selected at random students were not in direct competition with each other, thus minimizing the chances that the paradigm was inducing a competitive mind-set. Further, the terms *cooperation*,

competition, and *opponent* were expressly excluded from the instructions. Finally, students were told that at some point during but not necessarily at the midpoint of the game, there would be a break during which students would be asked to complete a written task.

The game then began. In truth, students did not actually play with each other, but rather played with a programmed tit-for-tat strategy that chose "A" (the cooperative choice) on the first trial, and thereafter chose what the student had chosen on the previous trial (Komorita, Hilty, & Parks, 1991). After the 15th trial, students were told that it was now time to complete the written task, titled "Performance Description." Students were instructed to think about the choices they had just made, outcomes they had received, and some alternatives that were worse or better than what actually happened. They were asked to write down as many such alternatives as they could think of. After they had generated the list of alternatives, they were told to place a plus sign next to each that described a better alternative and a minus sign next to each that described a worse alternative. This method of eliciting and evaluating thoughts is commonly used in counterfactual thought research (cf. Roese & Olson, 1995; Sanna, 1996).

After everyone had completed the thought-generation task, 15 more trials were conducted. After the 30th trial, students were informed that the game was over. They were then given a short questionnaire assessing their perceptions of the game and the extent to which they could predict the other person's choices. The former open-ended question was included to check for any suspicions students might have had about the deception. The latter question, which used a 9-point rating scale (1 = *not much*, 9 = *greatly*), was included because Komorita et al. (1991) reported some evidence that perceived predictability of the opponent influences own rate of cooperation, especially when close attention is being paid to the sequence of outcomes. Because we were asking students to think carefully about the sequence just completed, it seemed important to rule out predictability as an alternate explanation for the findings. After completing the questionnaire, students were thanked, debriefed (including a detailed discussion of the deception), and dismissed.

Results

Two students were eliminated from the data set because they failed to complete the thought-generation task correctly. No students indicated suspicion that they were actually interacting with the experimenter. Thus, all analyses are based on 78 data points.

For each student, we calculated pre- and post-thought-task rates of cooperation by dividing the number of cooperative choices before task and the number after task by 15. The posttask rate was used as the criterion in a hierarchical regression analysis. The predictors, in the order in which they were entered, were (a) pretask rate of cooperation, (b) predictability rating, (c) number of downward counterfactuals generated, and (d) number of upward counterfactuals generated. The first two were treated as covariates; the latter two were uncorrelated ($r = .06$) and so were entered in a random order. The ΔR^2 analyses for each model are shown in Table 1. It can be seen that the best overall model is the one with all predictors included. The weights for the best model are shown in Table 2. Not surprisingly, past cooperation is a predictor of future cooperation, and consistent with Komorita et al. (1991), predictability is positively related to future cooperation—the more predictable other seems to be, the more cooperative one is. The important results are those pertaining to the generation of counterfactual thoughts. Axelrod's (1984) and Camerer and Ho's (1998) models both imply that the number of upward counterfactuals generated should be positively related to future cooperation, and this is indeed the case. The more one thinks about how "it could have been better," the more cooperative one is in the future.

		OTHER PERSON	
		A	B
YOU	A	7 \ 7	0 \ 12
	B	12 \ 0	1 \ 1

Figure 1. A two-person prisoner's dilemma matrix.

Table 1
Hierarchical Regression Results—Study 1

Step	Predictor added	Model <i>F</i>	<i>R</i> ²	ΔR^2	Error <i>df</i>
1	Pretask cooperation	20.60	.21	.21	76
2	Predictability	15.22	.29	.08	75
3	Downward CF	14.05	.36	.07	74
4	Upward CF	13.56	.43	.06	73

Note. All ΔR^2 are significant at $\alpha = .01$. CF = counterfactual.

Note, however, that there is a negative relationship between the number of downward counterfactual thoughts generated and future cooperation. Thus, emphasizing how “it could have been worse” induces less cooperation in the future.

We have already noted that the two types of thoughts are not correlated. That they both independently affect future cooperation suggests that they act as counterbalances to each other. That is, the tendency to be cooperative that is encouraged by upward counterfactuals gets inhibited when one starts generating downward counterfactuals and vice versa. In support of this, we find that the average rate of postthought cooperation among those who generate only downward counterfactuals is significantly lower ($M = .23$, $n = 24$) than it is among those who generate at least one upward counterfactual ($M = .39$, $n = 54$; $z = -2.54$, Wilcoxon *W*; $p = .01$). Similarly, those who generate only upward counterfactuals are much more cooperative ($M = .53$, $n = 15$) than those who generate at least one downward counterfactual ($M = .30$, $n = 63$; $z = -2.66$, Wilcoxon *W*; $p < .01$). Note also the small number of participants (19%) who thought exclusively of how things could be better. Making at least one downward comparison was common within our sample.

Finally, although not directly related to our hypotheses, it was of interest to examine the content of the generated thoughts. Such analyses are often conducted in studies of counterfactual thinking (e.g., Sanna, 1996), because they can shed light on the motivations underlying the behavior of interest. The content of the statements suggested that participants fully understood the game and how it could be effectively played. Examples of upward counterfactual statements are “I would have more points now if we had both been choosing A,” and “If only I hadn’t been choosing B all the time, I’d probably be doing better.” Examples of downward counterfactuals are “Other guy chooses B if I keep choosing A and I don’t get anything,” and “At least I have some points.”

Discussion

Our data suggest that Axelrod (1984) and Camerer and Ho (1998, 1999a, 1999b) were accurate when they predicted that thinking about better states of affairs induces more frequent cooperation. However, our data also indicate that the concurrent presence of downward thoughts acts as a brake on the development of cooperation, and thinking only of better situations does not occur very often. Ruminating on the past thus does affect cooperation but not always in a desirable manner.

Study 2

The purpose of Study 2 was to examine the interplay between individual differences, counterfactual thought, and cooperation. It

is known that rate of cooperative choice as well as the number and type of counterfactual thoughts can be predicted from knowledge of certain individual-difference variables. In particular, we were interested in social value orientation, which predicts cooperation well. There is reason to believe that it might also predict the nature of counterfactual thoughts. Hence, our primary goal was to demonstrate that social value orientation, content of retrospection, and rate of cooperation relate to each other.

Social Value Orientation

Social value orientation refers to a person’s tendency to choose so as to maximize certain outcomes in an interdependent situation. Using Van Lange’s (1999) model, a *prosocial* individual desires to simultaneously maximize joint gain and equality in outcomes, an *individualist* is interested only in maximizing own gain, and a *competitor* wants to maximize relative gain (i.e., the difference between own and other outcome). This list is not exhaustive—see Komorita and Parks (1996) for examples of other orientations—but the vast majority of individuals can be described by one of these three categories.

Social value orientation is well established as an influence on rate of cooperative choice in prisoner’s-dilemma-type games (Komorita & Parks, 1996) as well as other types of mixed-motive situations (e.g., de Dreu, Giebels, & Van de Vliert, 1998; Parks, 1994). It is important to note for our purposes that value orientation seems also to affect cognitive processes that relate to the interdependent task. For example, research on the might versus morality effect has suggested that prosocials view cooperation as a moral issue (i.e., cooperation is the “right” thing to do), but competitors view it as a power issue (i.e., prosocials are weak and capitulative, competitors are strong and in control; Liebrand, Jansen, Rijken, & Suhre, 1986). That competitors think specifically in terms of power has not always been supported (Beggan, Messick, & Allison, 1988, Study 2; Sattler & Kerr, 1991), but even these authors largely reject the notion that prosocials and competitors see the question of cooperation in the same way. Relatedly, Roch, Samuelson, and colleagues (Roch, Lane, Samuelson, Allison, & Dent, 2000; Roch & Samuelson, 1997) have shown that social value orientation can influence how participants in a resource dilemma think about the resource. In particular, Roch et al. (2000) found that under conditions of low cognitive load, noncooperators tend to think primarily of how to deviate from equal distribution of the resource, whereas prosocials strike more of a balance between thoughts of deviation and thoughts of maintaining equality (see also de Dreu & Boles, 1998). Clearly, then, it is

Table 2
Regression Weights—Model 4, Study 1

Predictor	β	<i>t</i> (73)
Pretask cooperation	.39	4.17***
Predictability	.24	2.69**
Downward CF	-.30	-3.16***
Upward CF	.26	2.84***

Note. CF = counterfactual.
** $p = .01$. *** $p < .01$.

reasonable to expect differences in cognitive processing among the primary value orientations.

Research on counterfactual thinking has shown that particular individual differences can predict the number and content of retrospective thoughts. Specifically, defensive pessimism seems to influence counterfactual thought. Optimists tend to engage in downward simulation, but defensive pessimists emphasize how things could have been better (Sanna, 1996, 1998). A similar type of logic can be applied to competitors versus prosocials, especially if we consider the issue from the transformational standpoint suggested by McClintock and Liebrand (1988). From this perspective, the important point for predicting behavior is an understanding of how the person perceives the likelihood of a particular outcome occurring given expectations for how the other person will behave. With regard to competitors, they typically assume that others will be similarly competitive, and everyone will receive low payoffs (Kelley & Stahelski, 1970; but for an important exception, see Van Lange & Liebrand, 1991). Thus, when the competitor subjectively transforms the situation, only two states of affairs are considered realistic: competing when everyone else is competing or cooperating when everyone else is competing (McClintock & Liebrand, 1988). The former produces better outcomes than the latter (1 point versus no points in our Figure 1 example), so it follows that competitors will always compete. If we then ask competitors to reflect on past interactions, it should be the case that they consider only one alternate scenario: cooperating when others do not. Because this would have produced poorer outcomes, the competitors should conclude that things could have been worse than they are. Prosocials see the long-term value of sustained mutual cooperation. However, they must also have the expectation that others will be similarly cooperative before they themselves make cooperative choices (Pruitt & Kimmel, 1977; Yamagishi, 1986). Prosocials generally begin with a cooperative choice and continue to make such choices until others fail to act cooperatively, at which point they begin making uncooperative choices (Kelley & Stahelski, 1970; McClintock & Liebrand, 1988; Van Lange, 1999). Thus, unless the prosocial has been paired with other perfectly cooperative individuals, he or she will be able to imagine and should focus on how things could have been better.

Finally, individualists act at times like prosocials and at times like competitors (McClintock & Liebrand, 1988). Because their emphasis is solely on personal accumulation, they should be able to imagine both how they could be gaining more and how they could be gaining less.

Hypotheses

Given the rationale developed above, we expect to see different patterns of generation of upward and downward counterfactuals across prosocials, individualists, and competitors. Specifically, prosocials should generate the greatest number of upward counterfactuals and least number of downward counterfactuals, competitors should do exactly the reverse, and individualists should generate moderate amounts of both types of statements.

Participants

The participants were 90 students enrolled in an introductory psychology course. Participation was in partial fulfillment of a course requirement.

Method

The same payoff matrix and apparatus as in Study 1 were used. The only change was that before being introduced to the game, students completed the short-form Triple-Dominance Measure (Van Lange & Kuhlman, 1994). The measure consists of nine items. For each, the respondent must choose from among three different distributions of points to self and an anonymous other person. One distribution maximizes joint gain (e.g., 550 to self, 550 to other), one maximizes relative gain (e.g., 400 to self, 50 to other), and one maximizes personal gain (e.g., 600 to self, 350 to other). Those with a prosocial orientation should favor the first option, competitors the second, and individualists the last. An orientation is assigned if the person selects the option consistent with the same orientation at least six times. Those who fail to meet this criterion are considered unclassifiable and discarded. After administration and collection of this questionnaire, the procedure was the same as in Study 1.

Results and Discussion

Fifteen students (17%) failed to meet the criterion of six value-consistent choices and were discarded. This rate of attrition is consistent with past studies of social values (see Komorita & Parks, 1996). As in Study 1, no students indicated suspicion with the study. Thus, 75 students were available for analysis: 30 individualists, 26 prosocials, and 19 competitors.

Postthought rate of cooperation was again used as the dependent measure, and a hierarchical regression analysis was conducted with the following steps: (a) prethought cooperation; (b) predictability of other; (c) number of downward counterfactuals; (d) number of upward counterfactuals; and (e) value orientation, coded as prosocial = 1, individualist = 2, competitor = 3. It is interesting to note that in this data set the two counterfactual variables were correlated ($r = -.52, p < .01$).³ They were entered in the order (downward, upward) to mimic the Study 1 analysis.

The ΔR^2 analysis, the results of which are shown in Table 3, indicates that the best-fitting model includes prethought cooperation and the two counterfactual variables. Table 4 contains the regression weights associated with this model. Consistent with Study 1, number of downward counterfactuals was inversely related and number of upward counterfactuals was directly related to rate of cooperative choice. Unlike Study 1, predictability of other was unrelated to rate of cooperation. It is thus unclear to what extent predictability is a true nuisance variable in the study of cooperative choice.

Our major hypothesis for this study was that social value orientation can predict the type of counterfactual thoughts generated. Because the numbers of upward and downward counterfactuals were correlated, we tested for an effect of orientation on type of thought generated by conducting a multivariate regression, with the numbers of each type of thought as the criteria, and value orientation as the predictor. There was a significant effect (Wilks's $\lambda = 0.86$), $F(2, 72) = 5.83, p < .01$. Breaking the effect down into

³ Research on counterfactual thinking has not addressed whether upward and downward counterfactual statements should or should not correlate. One can predict theoretically that they should not correlate, because the two types of statements are presumed to serve different functions, but this has not been empirically established. Hence, we simply note that our correlations were of different magnitudes and refrain from identifying one as unusual.

Table 3
Hierarchical Regression Results—Study 2

Step	Predictor added	Model <i>F</i>	<i>R</i> ²	Δ <i>R</i> ²	Error <i>df</i>
1	Pretask cooperation	45.13	.38	.38	73
2	Predictability	22.34	.38	.00	72
3	Downward CF	47.28	.67	.28	71
4	Upward CF	39.71	.70	.03	70
5	Social value orientation	31.44	.70	.00	69

Note. Δ*R*² for Steps 1, 3, and 4 are significant at α = .01. CF = counterfactual.

its univariate components, we see that value orientation predicts both number of upward, $F(1, 73) = 6.03, p < .01$, and number of downward, $F(1, 73) = 5.86, p < .01$, counterfactuals. Cell means for each type are shown in Table 5. There is a neat relationship between number and orientation, and it is as predicted: Prosocials generated the greatest number of upward counterfactuals and the least number of downward counterfactuals, competitors did the reverse, and individualists generated the same amount of each type. Post hoc Tukey comparisons (using the harmonic mean of the cell counts) of each set of means showed that, for both types of counterfactual thought, prosocials and competitors differed significantly. We also checked to see if social value orientation can predict the total number of counterfactual thoughts generated (cf. Sanna, 1996, 1998) by treating the total as the criterion and value orientation as the predictor. The model was not significant ($F = 0.24$). The same number of thoughts was generated regardless of orientation ($M = 2.32$ for competitors, 2.31 for prosocials, and 2.17 for individualists).

It is surprising that value orientation was not a significant predictor of postthought cooperation. This can be explained by either value orientation being truly unrelated to rate of postthought cooperation or value orientation contributing nothing over and above the number of counterfactual thoughts generated. These explanations can be tested by means of a mediational analysis (Baron & Kenny, 1986). This involves deriving three regression equations: In Equation 1, the independent variable is predictor and mediator is criterion; in Equation 2, the independent variable is predictor and dependent variable is criterion; and in Equation 3, the independent variable and mediator are predictors, entered simultaneously, and dependent variable is criterion. If mediation is occurring, Equations 1 and 2 should both be significant, but only the mediator should be significant in Equation 3. Because we are proposing that two variables, numbers of upward and downward counterfactual thoughts, act together as a mediator, and because

Table 4
Regression Weights—Model 4, Study 2

Predictor	β	<i>t</i> (70)
Pretask cooperation	.43	6.14***
Predictability	.02	0.33
Downward CF	−.46	−5.72***
Upward CF	.20	2.51***

Note. CF = counterfactual.
*** $p < .01$.

Table 5
Mean Number of Each Type of Counterfactual Statement Generated by Persons With Each Type of Social Value Orientation

Social value	No. upward	No. downward
Prosocial	1.46	0.85
Individualist	1.00	1.17
Competitor	0.74	1.58

Note. Within each type of counterfactual thought, prosocials and competitors differ significantly (Tukey's honestly significant difference using harmonic mean of cell count).

those two variables were correlated in this data set, the analysis was modified such that Equation 1 was derived by means of a multivariate regression model.

Table 6 presents the weights associated with the analysis. The multivariate portion of Equation 1 proved significant, Wilks's $\lambda = 0.86, F(2, 72) = 5.83, p < .01$. The follow-up univariate regression analyses showed that social value orientation predicted both types of counterfactual statements. Equation 2 revealed that social value orientation was also predictive of postthought rate of cooperation. Finally, when both types of counterfactual statements and social value orientation were entered at once as predictors of cooperation in Equation 3, the latter failed to achieve significance. This pattern is exactly as it should be if mediation is occurring: Social value orientation predicts number of each type of counterfactual thought as well as postthought rate of cooperation, but when value orientation and thought types are considered together as predictors of cooperation, social value orientation is no longer predictive.

Next, as in Study 1, we examined the rates of cooperation among those who generated just upward counterfactuals or just downward counterfactuals. As in Study 1, people who generated only downward counterfactuals were significantly less cooperative ($M = .26, n = 20$) than were people who generated at least one upward counterfactual thought ($M = .55, n = 60; z = -4.00$, Wilcoxon *W*; $p < .01$). The effect associated with downward counterfactuals also replicated: Those who generated only upward counterfactuals were much more cooperative ($M = .76, n = 21$)

Table 6
Regression Weights for Test of the Hypothesis That Counterfactual Thoughts Mediate the SVO–Cooperation Link

Step	Criterion	Predictor	β	<i>t</i>
1	UCF	SVO	−.34	−3.09*** ^a
		DCF	.31	2.76*** ^a
2	Cooperation	SVO	−.28	−3.02*** ^b
		UCF	.20	2.07* ^c
3	Cooperation	DCF	−.57	−6.07*** ^c
		SVO	−.11	−1.28 ^c

Note. In Step 1, a multivariate regression was first performed, with upward counterfactual (UCF) and downward counterfactual (DCF) as the criteria and social value orientation (SVO) as the predictor. For that analysis, Wilks's $\lambda = 0.86, F(2, 72) = 5.83, p < .01$.

^a $df = 73$. ^b $df = 104$. ^c $df = 71$.

* $p < .05$. *** $p < .01$.

than were those who made at least one downward comparison ($M = .37$, $n = 59$; $z = -4.88$, Wilcoxon W; $p < .01$).

Finally, also following Study 1, we informally examined the content of the thought statements. As in Study 1, the statements discussed the different strategies and payoff combinations inherent in the game—a sample upward counterfactual was “If only we could get together and both choose A more, we would both do really well,” and a sample downward counterfactual was “I could have 0 points right now, at least I don’t.”

General Discussion

The primary purpose of this study was to investigate whether retrospection about a social dilemma can influence subsequent cooperation. We were motivated by both Axelrod’s (1984) and Camerer and Ho’s (1998) arguments for the importance of retrospection in the cooperation process. Across two studies, we saw that thinking in terms of how things could have been worse inhibits future cooperation, and thinking about how things could have been better enhances future cooperation, though the presence of both types of thoughts seems to temper the effects of each. Our second study additionally showed a connection between social value orientation and type of counterfactual thought, in that prosocials generate primarily upward counterfactuals, competitors primarily downward counterfactuals, and individualists equal numbers of both. We would thus say that Axelrod and Camerer and Ho were correct in advocating contemplation of better outcomes as a means of fostering cooperation, but note that people seem to rarely think only of better outcomes, and our data suggest that cooperation is enhanced only when “could have been worse” thoughts are avoided.

There are many implications associated with these results, for both theory on cooperative choice and theory on counterfactual thinking. For the study of cooperation, first and foremost, the results provide a rationale for testing for the mood–cooperation effects suggested in the introduction of this article, and we have initiated such tests. Relatedly, research on defensive pessimism has shown that defensive pessimists, who generate primarily upward thoughts, seem to use such scenarios as motivators, at least when generated prefactorially; that is, defensive pessimists anticipate that things will go poorly, imagine how things might turn out better, and then elevate their performance to try to achieve the imagined better outcomes (e.g., Norem & Cantor, 1986). Upward counterfactuals may be similarly motivating for future cooperation. We saw in both studies that people who generate strictly upward counterfactuals were highly cooperative after the thought-generation process. Perhaps imagining how things could be better, without simultaneously thinking about how things could have been worse, encourages people to try to attain the better outcomes.

Our data also add to the existing literature on counterfactual thinking by providing further evidence that mentally simulating alternative outcomes can influence future behaviors (Roese & Olson, 1995; Sanna, 1996). We found that it was not only upward counterfactuals that were associated with future choices, but downward ones were also uniquely associated with people’s choice behaviors. There are few studies that have actually linked counterfactuals to future behaviors (for a review, see Roese, 1997), and so our research also takes another important step in that regard.

There are many variables that might influence the counterfactual thought–cooperation link. An important one is the opponent’s strategy. For example, consider a person who is in a cooperative equilibrium with other (i.e., both consistently choose “A”). During reflection, the person may realize that he or she could be doing even better if other could be exploited, that is, if other will continue to choose “A” while the person begins choosing “B.” This would represent an instance of an upward counterfactual leading to a decrease in cooperation. Such exploitation was not possible in our studies because the programmed other played tit for tat, but there are many strategies that are more lenient and can be exploited with some success. Hence, the motivating properties of upward counterfactual thoughts may well interact with exploitability of other. Providing some support for this idea is the work of Van Lange and Visser (1999), who showed that competitors prefer to end interactions with others who use tit for tat. Their explanation is that competitors quickly realize that they cannot achieve consistent exploitation of other. From our perspective, it could be argued that competitors generate the upward counterfactual, “I could be doing much better if I could exploit somebody,” and decide to seek a relationship in which this can occur.

Another variable of potential interest is the motivation for behavior. Research has shown that, although cooperative choice is typically motivated by a desire for all to benefit, noncooperation is sometimes motivated by greed and sometimes by fear, either of receiving nothing for one’s efforts or of being exploited (e.g., Kerr, 1983; Parks & Hulbert, 1995; Rapoport & Eshed-Levy, 1989). Relevant to our purposes, Camac (1992) has documented relationships between social value orientation and preferences for information about the fear, greed, and mutual benefits inherent within a dilemma. Specifically, he found that noncooperators were most interested in greed-related information, and prosocials were most interested in information that contrasted mutual-cooperation outcomes against mutual-noncooperation outcomes. (The two groups were equal on their preference for fear-related information.) Camac attributed these differing preferences to differences in schematic processing, but it is also possible that they could be driven by counterfactual thinking. A desire to know how mutual noncooperation compares with mutual cooperation could be driven by upward counterfactual thoughts (e.g., “We’re getting 1 each but we could probably do better if we were all cooperating”). However, the desire for greed information could also result from upward counterfactual thinking—as we noted in the previous paragraph, noncooperators who are not confronted with a tit-for-tat–using partner (and in Camac’s study, they were not) can realistically think about achieving the best exploitative payoff. Here the same type of thought is being generated, but for different reasons and to different ends. Our suppositions about opponent strategy and underlying motivation suggest that the type of behavior that results from a counterfactual thought may be situation specific. It will be important to determine when to expect what type of behavior.

Conclusion: Counterfactual Thoughts, Social Value Orientation, and Cooperation

Potentially one of our most important findings is that in Study 2, social value orientation added nothing to the predictive ability of the regression model after the two types of counterfactual thoughts were partialled out. This suggests that a considerable portion of the

differences between prosocials, individualists, and competitors can be attributed to the way in which they think about unrealized outcomes. Prosocials seem to focus on how things could be better, competitors focus on how things could be worse, and individualists examine both perspectives. That people with different social value orientations would focus on different aspects of the same experience is consistent with work showing that these people highlight different features of the same opponent-choice strategy (Parks & Rumble, 2001) and also fits within current theory on the development of social value orientations.

Van Lange, Otten, de Bruin, and Joireman (1997) have argued that although value orientation may have some genetic basis (see, e.g., Rushton, Fulker, Neale, Nias, & Eysenck, 1986), above and beyond this, differences in value orientation are produced by the social interaction experiences (Kelley, 1997) that accrue as one is exposed to situations of interdependence early in life. These experiences, which are a function of both the nature of the interdependence (e.g., mixed motive, zero sum, high correspondence of outcomes) and the actions of the others with whom one is involved, are assumed to affect how one transforms an objective outcome set (or *given matrix*) into a subjective outcome set (*effective matrix*; Kelley & Thibaut, 1978). Thus, from Van Lange et al.'s perspective, a social value orientation is basically a tendency to transform a given matrix in a particular way. Recall that our predictions for Study 2 were derived from McClintock and Liebrand's (1988) transformational perspective. Hence, our results provide indirect support for Van Lange et al.'s notion that people with different social value orientations transform given matrices in different ways. Our results further suggest that, as a consequence of these differing transformations, individuals with different social value orientations evaluate the past in different ways, and these various evaluations in turn influence future choice strategy. This inference is supported by the results of the mediational analysis in Study 2, which suggests some type of causal linkage between social value orientation, generation of counterfactual thoughts, and cooperation.

We can combine our results with those of Van Lange et al. (1997) to create a tentative model of how social value orientation develops and affects behavior. To wit, as a result of social interaction experiences that occur during early life (and possibly across the life span), and possibly with some genetic influence, the individual develops a tendency to transform an objective set of outcomes in one particular way. This preferred transformation influences which aspects of the interdependence the person attends to when he or she reflects on the relationship (and interdependence theory predicts that such reflection occurs regularly; Thibaut & Kelley, 1959). The nature of the reflection in turn affects the behavior of the person in future interdependent interactions.

In sum, then, our two studies make contributions to both the understanding of the dynamics of counterfactual thought and, especially, of the nature of social value orientation. More generally, our studies also demonstrate the value of examining cognitive influences on cooperative choice, an approach that is rapidly gaining popularity within the social dilemmas field (e.g., de Bruin & Van Lange, 1999; de Dreu & McCusker, 1997; de Dreu, Yzerbyt, & Leyens, 1995; Parks, Sanna, & Berel, 2001; Samuelson & Allison, 1994).

References

- Anderson, C. M., & Camerer, C. (2000). Experience-weighted attraction learning in sender-receiver signaling games. *Economic Theory*, *16*, 689-718.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.
- Beggan, J. K., Messick, D. M., & Allison, S. T. (1988). Social values and egocentric bias: Two tests of the might over morality hypothesis. *Journal of Personality and Social Psychology*, *55*, 606-611.
- Camac, C. (1992). Information preferences in a two-person social dilemma. In W. B. G. Liebrand, D. M. Messick, & H. A. M. Wilke (Eds.), *Social dilemmas* (pp. 147-161). New York: Pergamon Press.
- Camerer, C., & Ho, T.-H. (1998). Experience-weighted attraction learning in coordination games: Probability rules, heterogeneity, and time-variation. *Journal of Mathematical Psychology*, *42*, 305-326.
- Camerer, C., & Ho, T.-H. (1999a). Experience-weighted attraction learning in games: Estimates from weak-link games. In D. V. Budescu, I. Erev, & R. Zwick (Eds.), *Games and human behavior* (pp. 31-51). Mahwah, NJ: Erlbaum.
- Camerer, C., & Ho, T.-H. (1999b). Experience-weighted attraction learning in normal form games. *Econometrica*, *67*, 827-874.
- De Bruin, E. N. M., & Van Lange, P. A. M. (1999). Impression formation and cooperative behavior. *European Journal of Social Psychology*, *29*, 305-328.
- De Dreu, C. K. W., & Boles, T. L. (1998). Share and share alike or winner take all? The influence of social value orientation upon choice and recall of negotiation heuristics. *Organizational Behavior and Human Decision Processes*, *76*, 253-276.
- De Dreu, C. K. W., Giebels, E., & Van de Vliert, E. (1998). Social motives and trust in integrative negotiation: The disruptive effects of punitive capability. *Journal of Applied Psychology*, *83*, 408-422.
- De Dreu, C. K. W., & McCusker, C. (1997). Gain-loss frames and cooperation in two-person social dilemmas: A transformational analysis. *Journal of Personality and Social Psychology*, *72*, 1093-1106.
- De Dreu, C. K. W., Yzerbyt, V. Y., & Leyens, J. P. (1995). Dilution of stereotype-based cooperation in mixed-motive interdependence. *Journal of Experimental Social Psychology*, *31*, 575-593.
- Hertel, G. (1999). Mood effects in social dilemmas: What we know so far. In M. Foddy, M. Smithson, S. Schneider, & M. Hogg (Eds.), *Resolving social dilemmas: Dynamic, structural, and intergroup aspects* (pp. 227-244). New York: Psychology Press.
- Kelley, H. H. (1997). The "stimulus field" for interpersonal phenomena: The source of language and thought about interpersonal events. *Personality and Social Psychology Review*, *1*, 140-169.
- Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, *16*, 66-91.
- Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations*. New York: Wiley.
- Kerr, N. L. (1983). Motivational losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology*, *45*, 819-828.
- Komorita, S. S., Hilty, J. A., & Parks, C. D. (1991). Reciprocity and cooperation in social dilemmas. *Journal of Conflict Resolution*, *35*, 494-518.
- Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology*, *46*, 183-207.
- Komorita, S. S., & Parks, C. D. (1996). *Social dilemmas*. Boulder, CO: Westview Press.
- Liebrand, W. B. G., Jansen, R. W. T. L., Rijken, V. M., & Suhre, C. J. M. (1986). Might over morality: Social values and the perception of other

- players in experimental games. *Journal of Experimental Social Psychology*, 22, 203–215.
- Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology*, 29, 87–109.
- Martin, L. L., Ward, W., Achee, J. W., & Wyer, R. S. (1993). Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, 64, 317–326.
- McClintock, C. G., & Liebrand, W. B. G. (1988). Role of interdependence structure, individual value orientation, and another's strategy in social decision making: A transformational analysis. *Journal of Personality and Social Psychology*, 55, 396–409.
- Norem, J. K., & Cantor, N. (1986). Defensive pessimism: Harnessing anxiety as motivation. *Journal of Personality and Social Psychology*, 51, 1208–1217.
- Parks, C. D. (1994). The predictive ability of social values in resource dilemmas and public goods games. *Personality and Social Psychology Bulletin*, 20, 431–438.
- Parks, C. D., & Hulbert, L. G. (1995). High and low trusters' responses to fear in a payoff matrix. *Journal of Conflict Resolution*, 39, 718–730.
- Parks, C. D., & Rumble, A. C. (2001). Elements of reciprocity and social value orientation. *Personality and Social Psychology Bulletin*, 27, 1301–1309.
- Parks, C. D., Rumble, A. C., & Posey, D. C. (2002). The effects of envy on reciprocation in a social dilemma. *Personality and Social Psychology Bulletin*, 28, 509–520.
- Parks, C. D., Sanna, L. J., & Berel, S. R. (2001). Actions of similar others as inducements to cooperate in social dilemmas. *Personality and Social Psychology Bulletin*, 27, 345–354.
- Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28, 363–392.
- Rapoport, A., & Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gypped. *Organizational Behavior and Human Decision Processes*, 44, 325–344.
- Roch, S. G., Lane, J. A. S., Samuelson, C. D., Allison, S. T., & Dent, J. L. (2000). Cognitive load and the equality heuristic: A two-stage model of resource overconsumption in small groups. *Organizational Behavior and Human Decision Processes*, 84, 185–212.
- Roch, S. G., & Samuelson, C. D. (1997). Effects of environmental uncertainty and social value orientation in resource dilemmas. *Organizational Behavior and Human Decision Processes*, 70, 221–235.
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121, 133–148.
- Roese, N. J., & Olson, J. M. (1995). Outcome controllability and counterfactual thinking. *Personality and Social Psychology Bulletin*, 21, 620–628.
- Rushton, J. P., Fulker, D. W., Neale, M. C., Nias, D. K. B., & Eysenck, H. J. (1986). Altruism and aggression: The heritability of individual differences. *Journal of Personality and Social Psychology*, 50, 1192–1198.
- Samuelson, C. D., & Allison, S. T. (1994). Cognitive factors affecting the use of social decision heuristics in resource sharing tasks. *Organizational Behavior and Human Decision Processes*, 58, 1–27.
- Sanna, L. J. (1996). Defensive pessimism, optimism, and simulating alternatives: Some up and downs of prefactual and counterfactual thinking. *Journal of Personality and Social Psychology*, 71, 1020–1036.
- Sanna, L. J. (1998). Defensive pessimism and optimism: The bitter-sweet influence of mood on performance and prefactual and counterfactual thinking. *Cognition and Emotion*, 12, 635–665.
- Sanna, L. J. (2000). Mental simulation, affect, and personality: A conceptual framework. *Current Directions in Psychological Science*, 9, 168–173.
- Sanna, L. J., Stocker, S. L., & Clarke, J. A. (2003). Rumination, imagination, and personality: Specters of the past and future in the present. In E. C. Chang & L. J. Sanna (Eds.), *Virtue, vice, and personality: The complexity of behavior* (pp. 105–124). Washington, DC: American Psychological Association.
- Sanna, L. J., Turley-Ames, K. J., & Meier, S. (1999). Mood, self-esteem, and simulated alternatives: Thought-provoking affective influences on counterfactual direction. *Journal of Personality and Social Psychology*, 76, 543–558.
- Sattler, D. N., & Kerr, N. L. (1991). Might versus morality explained: Motivational and cognitive bases for social motives. *Journal of Personality and Social Psychology*, 60, 756–765.
- Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.
- 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.
- Van Lange, P. A. M., & Kuhlman, D. M. (1994). Social value orientations and impressions of a partner's honesty and intelligence: A test of the might versus morality effect. *Journal of Personality and Social Psychology*, 67, 126–141.
- Van Lange, P. A. M., & Liebrand, W. B. G. (1991). The influence of other's morality and own social value orientation on cooperation in the Netherlands and the USA. *International Journal of Psychology*, 26, 429–447.
- Van Lange, P. A. M., Otten, W., de Bruin, E. M. N., & Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology*, 73, 733–746.
- Van Lange, P. A. M., & Visser, K. (1999). Locomotion in social dilemmas: How we adapt to cooperative, tit-for-tat, and noncooperative partners. *Journal of Personality and Social Psychology*, 77, 762–773.
- Yamagishi, T. (1986). The structural goal/expectation theory of cooperation in social dilemmas. *Advances in Group Processes*, 3, 51–87.

Received December 18, 2001

Revision received June 20, 2002

Accepted June 21, 2002 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://watson.apa.org/notify/> and you will be notified by e-mail when issues of interest to you become available!