

University of Groningen

## Value orientation and conformity. A study Using three types of social dilemma games

Vogel, Rob; Wilke, Henk A.M.; Leibrand, Wim B.G.; Wolters, Fred J.M.

*Published in:*  
Journal of Conflict Resolution

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
1986

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Vogel, R., Wilke, H. A. M., Leibrand, W. B. G., & Wolters, F. J. M. (1986). Value orientation and conformity. A study Using three types of social dilemma games. *Journal of Conflict Resolution*, 30(1), 76-98.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Value Orientation and Conformity

## A STUDY USING THREE TYPES OF SOCIAL DILEMMA GAMES

WIM B.G. LIEBRAND  
HENK A.M. WILKE  
ROB VOGEL  
FRED J.M. WOLTERS

*Department of Psychology  
University of Groningen*

---

---

Three different types of N-person social dilemma games were employed: the Prisoner's Dilemma (NPD), the Chicken Dilemma (NCD), and the Trust Dilemma (NTD). Subjects, who were classified a priori as either a cooperator (n = 58) or a defector (n = 68), participated in one of the social dilemma games before they received bogus feedback: they were told that the majority had chosen the defecting alternative D, or that the majority had chosen the cooperative alternative C. As predicted, (1) both before and after feedback, more defecting choices were made in the NPD than in the NCD, whereas in the NCD more defecting choices were made than in the NTD; (2) before and after feedback, defectors made more defecting choices than cooperators; (3) after majority D feedback more defecting choices were made than after majority C feedback. In addition, it appeared that in NPD and in NTD, defectors were especially sensitive to majority D feedback in that it facilitated their natural inclination to prefer D-choices. No support for Kelley and Stahelski's triangle hypothesis was observed.

---

---

Situations in which the private interest can be at odds with the public interest constitute an important class of societal problems. Undoubtedly, the most famous example is the Lloyd-Hardin metaphor of "the tragedy of the commons" (Hardin and Baden, 1977). Other examples

---

**AUTHORS' NOTE:** We wish to thank David M. Messick for his helpful comments on an earlier draft of the manuscript. Requests for reprints should be sent to Wim B.G. Liebrand, Department of Psychology, Grote Markt 31/32, 9712 HV Groningen, The Netherlands.

JOURNAL OF CONFLICT RESOLUTION, Vol. 30 No. 1, March 1986 77-97  
© 1986 Sage Publications, Inc.

are population control (Kahan, 1974), pollution control management (Dawes et al., 1974) and the use of commonly-owned resources (Hamburger, 1979).

All these so-called social dilemmas can be considered as binary situations in which two strategies are available: either choose alternative C (cooperate) in order to serve the public interest, or choose alternative D (defect), which serves the immediate private interest.

Formal definitions of social dilemmas are provided by Dawes (1975) and Liebrand (1983). Dawes (1975) defines a social dilemma as a situation in which: (1) each person has available a dominant strategy, in other words, one that yields the person the best payoff in all circumstances; and (2) in which the collective choice of dominating strategies results in a deficient outcome, that is, a result that is less preferred by all persons than the result which would have occurred if all had not chosen their dominating strategy.

Liebrand (1983) argues that a dominating strategy is not a necessary property for considering a situation a social dilemma. Relaxing the dominance assumption, Liebrand defines a social dilemma as a situation in which: (1) each person has available a strategy that yields the person the best payoff in at least one configuration of strategy choices, and that has a negative impact on the interests of the other persons involved (referred to as the D [defecting]-strategy) and in which (2) the choice of the D-strategy by all persons results in a deficient outcome. The relaxation of the dominance assumption makes the Liebrand definition broader than the Dawes definition that incorporates only the so-called Prisoner's Dilemma situations. In the present study the Liebrand definition of social dilemmas is employed.

Within the behavioral sciences, the strongest formal representation of social dilemma situations can be found in conceptual and empirical work on experimental games. In the last decade, a number of studies have been conducted within this social dilemma game paradigm (for reviews see Dawes, 1980; Messick and Brewer, 1983; Wilke et al., 1983).

This study extends prior research in that it uses three different types of social dilemma games (i.e., the Prisoner's Dilemma, the Chicken Dilemma and the Trust Dilemma). In addition, within and across these three games the effect of preexisting individual differences in social values on choice behavior, and the effect of the manipulation of majority feedback about others' choice behavior are assessed. In the following we will discuss each of these variables, and formulate three hypotheses.

### THREE TYPES OF N-PERSON SOCIAL DILEMMA GAMES

#### HYPOTHESIS 1

Starting from Rapoport and Guyer's (1966) taxonomy of two-person symmetric games, in which each player has a strict preference ordering of the four possible payoffs, Liebrand (1983) has demonstrated that three types of N-person games may be distinguished that meet his definition of social dilemmas. An N-person social dilemma can be a Prisoner's, a Chicken, or a Trust Dilemma. The three types of games are described below. The following notation is used. Each of N players has a choice between the D and the C strategy.

Let j players choose C, then  $D_j$  refers to the player's payoff for a D choice, and  $C_j$  refers to the player's payoff for a C choice.

A game is classified as an N-person Prisoner's Dilemma (NPD) if and only if

$$D_j > C_{j+1} \quad \text{for } j = 0, \dots, N-1 \quad [1]$$

$$C_N > D_0 \quad [2]$$

Condition 1 guarantees that choosing D is a dominant strategy. In the two other types of social dilemmas there is no dominant strategy available.

In an N-person Chicken Dilemma (NCD) the D choice yields the player the best payoff only if more than p other players choose C ( $0 \leq p < N-1$ ). Thus, games for which in addition to condition 2, conditions 3 and 4 hold are classified as NCD games

$$D_j > C_{j+1} \quad \text{for } j = p+1, \dots, N-1 \quad [3]$$

$$D_j \leq C_{j+1} \quad \text{for } j = 0, \dots, p \quad [4]$$

Finally, in an N-person Trust Dilemma (NTD) the D choice yields the player the best payoff if less than p other players choose C ( $0 \leq p < N-1$ ). Consequently, if condition 2 and conditions 5 and 6 hold, the game is classified as an NTD game

$$D_j > C_{j+1} \quad \text{for } j = 0, \dots, p \quad [5]$$

$$D_j \leq C_{j+1} \quad \text{for } j = p + 1, \dots, N - 1 \quad [6]$$

Until now empirical research has been focused primarily on the Prisoner's social dilemma (Kelley and Grzelak, 1972; Caldwell, 1976; Bonacich et al., 1976; Dawes et al., 1977; Dawes and Orbell, 1981; van de Kragt et al., 1983). Only a few studies employed the Chicken and Trust Dilemma (Meux, 1973; Tyszka and Grzelak, 1976). It is our opinion that the relative lack of interest for NCD and NTD situations is undesirable for two reasons.

First, as far as social dilemma research is concerned with real life dilemmas, the experimental games used to investigate choice behavior in social dilemma situations should be appropriate simulations of decisional structure underlying real life dilemmas (see Hamburger, 1979: pp. 247-248; Colman, 1982: p. 184). Liebrand (1983) describes three real-life situations (i.e., pollution, congestion, and hoarding), of which the underlying decisional structure is properly captured by an NPD, an NCD, and an NTD game, respectively. It follows that the ecological validity of social dilemma simulations is increased by investigating choice behavior in three types of social dilemmas.

Second, psychological or social factors that lead to individual cooperative choice behavior may not have the same effect in all three types of social dilemmas. In order to help develop solutions to real life dilemmas, social dilemma research could be used as a hypothesis testing device (see Dawes, 1980). In that case, besides NPD, also NCD and NTD games should be included in the empirical research on experimental games.

The payoff matrices of the three games used in this article are shown in Table 1. In constructing the games we tried in two ways to maximize the comparability of the games subject to the constraints 1 to 6. First, the payoffs for a C-choice were held constant. Second, the intersection of the payoff functions for NCD and NTD was fixated at the point at which about half of the subjects choose C.

In the following we will consider how the type of game affects choice behavior.

Based on Rapoport's (1966) index of cooperation for the  $2 \times 2$  PDG both Harris (1969, 1972) and Komorita (1976) proposed an index of cooperation for a broad range of experimental games. These indices assume that there are two basic motivational pressures guiding choice behavior: fear, that is, the desire to avoid the sucker's payoff; and greed, that is, the temptation to defect. Following Komorita et al.'s (1980) suggestion, in the present study we will use an extension of Harris's index

TABLE 1  
Payoff Matrices for the NPD, the NCD, and the NTD Games

Choice Configuration	Type of Game					
	NPD		NCD		NTD	
	Payoff for D	Payoff for C	Payoff for D	Payoff for C	Payoff for D	Payoff for C
0 D - 8 C	-	+8	-	+8	-	+8
1 D - 7C	+11	+6	+13	+6	+1.5	+6
1 D - 6C	+9	+4	+9	+4	+1	+4
3 D - 5C	+7	+2	+5	+2	+ .5	+2
4 D - 4C	+5	0	+1	0	0	0
5 D - 3C	+3	-2	-3	-2	-.5	-2
6 D - 2C	+1	-4	-7	-4	-1	-4
7 D - 1C	-1	-6	-11	-6	-1.5	-6
8 D - 0C	-3	-	-15	-	-2	-

of cooperation because the slopes of the payoff functions are not equal. Fear is estimated by:  $r_3 = (D_0 - C_1) / (D_{N-1} - C_1)$  and greed is estimated by:  $r_4 = (D_{N-1} - C_N) / (D_{N-1} - C_1)$ . The index of cooperation (K) is then defined as:  $K = 1 - w_1(r_3) - w_2(r_4)$  in which  $w_1$  and  $w_2$  represent the weights assigned to fear and greed, respectively.

Under the constraint that the weight for fear equals the weight for greed, the K-index for NPD = .65; for NCD:  $K = 1.21$ ; and for NTD:  $K = 1.33$ . However, as Alcock and Mansell (1977), and Komorita et al. (1980) have pointed out, the equality of the fear and greed components may be questioned. They suggest that the weight that has to be assigned to greed is greater than the weight for fear. Because such a modification does not affect the present rank-order of K-indices, we predict a rank-order among the games in number of D-choices to be expected. Hypothesis 1 therefore states that the NPD-game will evoke more D-choices than the NCD-game, whereas the NCD-game in its turn will yield more D-choices than the NTD-game. Finally, given the data of the present study, it is possible to compute the size of the relative weights which (at least in this study) have to be assigned to fear and greed.

#### VALUE ORIENTATION: HYPOTHESIS 2

It has been found repeatedly that value orientation, that is, a subject's preference for a particular distribution of outcomes to self and to others, affects choice behavior in a variety of experimental games (Messick and McClintock, 1968; McClintock, 1972; Kuhlman and Marshello, 1975; Bem and Lord, 1979; Liebrand, 1984; Liebrand and Van Run, 1985).

Usually four classes of value orientation are distinguished; altruism: the orientation to maximize others' outcomes; cooperation: maximize the sum of the own and other's outcomes; individualism: maximize own outcomes; and competition: maximize own relative advantage. However, in order to avoid a large number of null subclasses in the design of the present study, we will distinguish between only two classes. Subjects having an altruistic or a cooperative value orientation are classified as "cooperators," and those having an individualistic or a competitive value orientation are classified as "defectors."

Prior to their participating in either an NPD, an NCD, or an NTD, subject's value orientation was assessed by means of a decomposed games technique (Pruitt, 1967; Messick and McClintock, 1968; Griesinger and Livingston, 1973; Liebrand, 1984) to be described in more detail in the procedure section.

The first reason to include value orientation in the design of the present study was the possibility of generalizing the effect of value orientation for the three types of social dilemma settings. The second reason stems from the fact that previous research has not been directed to a possible difference between cooperators and defectors in their sensitivity to feedback about what the majority of the other players in an *n*-person game had chosen. Because there seems to be no *a priori* reason for cooperators to be more sensitive to such feedback than defectors, or vice versa, no hypothesis concerning this specific interaction was formulated.

Based on the above research findings, we do expect differences to occur in choice behavior between subjects classified *a priori* as cooperative or as defecting. Cooperative subjects are expected to choose in such a way that the collective payoff is maximized, that is they are expected to choose C in all three types of games. Defecting subjects on the other hand are expected to choose in such a way that the payoff to self is maximized. In NPD, payoff to self is maximized by choosing D. In NCD and NTD, the payoff to self-maximization depends upon the number of other subjects expected to choose C. However, both for NCD and NTD there are choice configurations that prescribe to choose D in case one wants to maximize the payoff to self. Hence, across the three types of games it is expected that defectors make more D-choices than cooperators (hypothesis 2).

### **CONFORMITY: HYPOTHESIS 3**

In the present study each subject participates on eight successive trials in either an NPD, an NCD, or an NTD. The subjects receive no

information about the choices of others except after the third trial. After the third trial, the subjects receive bogus feedback on collective choice behavior in the first, second, and third trial: they are being told either that a majority chose the cooperative strategy or that a majority chose the defecting strategy in each of the preceding trials. Such information about collective choice behavior may affect subjects' expectation about others' choice behavior and, in its turn, these expectations may influence their actual choice behavior.

Informal communication theory (Festinger, 1950) and social comparison (Festinger, 1954) suggest that individuals conform to the majority of their group, because this serves two functions. The informational function is that a majority may help an individual to select the best response in a particular situation; the normative function refers to the idea that an individual may wish to conform to the expectations of the other group members, because by doing so he or she expects to be rewarded by fellow group members.

There is abundant evidence that one is inclined to conform to the behavior of other group members, in a variety of stimulus situations (Allen, 1965; Kiesler, 1969; Shaw, 1976), including social dilemma situations (Klandermans, 1983; Messick et al., 1983; Schroeder et al., 1983).

In the light of the above evidence it seems justified to predict that in our three social dilemma situations, subjects who get to know that the majority of their group did select the cooperative strategy will act more cooperatively than subjects who hear that the majority of their group has opted for a defecting strategy. This prediction constitutes our hypothesis 3.

## METHOD

*Design.* The present study was divided into two parts: (1) choice behavior before feedback; and (2) choice behavior after bogus feedback. Type of game and value orientation constituted the two experimental factors in the (3 × 2) design of part one: subjects (cooperative or defecting) either played an NPD, an NCD, or an NTD game for three trials. The dependent variable was choice behavior before feedback. Expectations of subjects about other players' choice behavior were measured for explorative purposes only.

In part two, choice behavior was studied as a function of type of game (3 levels), value orientation (2 levels) and majority feedback (2 levels).



The feedback manipulation consisted of telling the subjects that the majority had chosen either C or D on each of the first three trials. The dependent variables in part two were the actual choice behavior of subjects and their expectations about others' choice behavior.

*Subjects.* Subjects, 144 volunteer students (age 18 to 30) from a Higher Institute of Technology in Enschede, The Netherlands, were randomly assigned to one of 18, 8-person groups. Subjects responded to an advertisement in the campus paper, in which it was announced that, depending on their performances, participants could earn:  $f7,50$  (about \$2.50) to  $f40,=$  (about \$13.30).

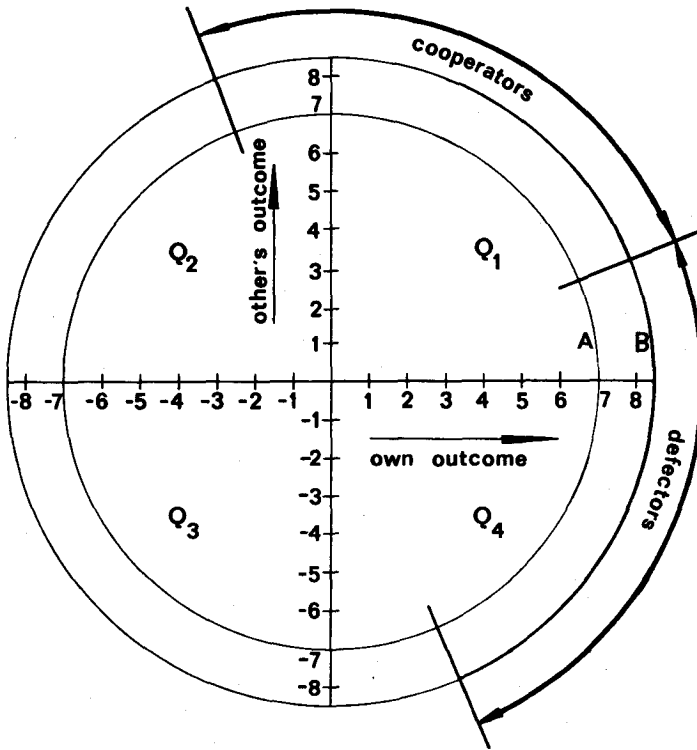
Eleven subjects were discarded because they indicated that they were suspicious of the bogus feedback. That is, when the experimenter asked them afterwards, in person, whether they fully believed the feedback after the third trial, they denied that. Most suspicion occurred in the NTD-majority defecting condition: eight subjects in this condition doubted the feedback.

In addition, as is described below in further detail, seven subjects were discarded because they could not be classified as defector or cooperators. This left 126 subjects, 116 males and 10 females; 41 in the NPD-, 46 in the NCD-, and 39 in the NTD-condition. The distribution of females over the 3 game conditions appeared to be random ( $\chi^2[2] = 4.2, p > .10$ ).

## PROCEDURE

The subjects were individually invited to the laboratory, where they took part in the experiment in groups of eight subjects. They were placed around a table ( $4 \times 4$  meter) behind small screens so that they could see one another without seeing their response sheets. During the experiment, subjects were not allowed to communicate with one another.

First, decomposed games were administered. The decomposed games technique employed is described in more detail in Liebrand (1984). It consists of making 32 choices between two own/other outcome combinations. Each own/other outcome combination is defined as a point on either Circle A (radius \$7.00) or on Circle B (radius \$8.50) in the outcome plane depicted in Figure 1. There were 16 equally spaced pairs of points on each of the circles, as each pair consisted of two adjacent points. An example of such a pair is the choice between \$3.30



NOTE: Q1 to Q4 = quadrant 1 to quadrant 4. Circle A radius = \$7.00, circle B radius = \$8.50

Figure 1: Own-Other Outcome Space

for self and \$7.90 for the other, versus \$6.00 for self and \$6.00 for the other. For each pair the subjects selected the point they preferred most.

Subjects were told that the other person in the decomposed games was another person whom they would never know, and they were told that the other player's choices would remain unknown, so as to avoid considerations of strategy. In the instructions, the structure of the decomposed games was thoroughly explained by means of two examples, but no advice was given on how to select the outcomes.

Adding up the chosen amounts separately for self and for other yields an estimate of the subject's value vector (Liebrand, 1984). Subjects were classified into one of the two classes of value orientation if the percentage of choices identical with a consistent choice pattern according to one of the two value orientations exceeded 60%. The average number of consistent choices for the classifiable subjects was

88%. If subjects were classifiable, all observed value vectors in between degree 112.5 and 22.5 were labelled cooperative, and all vectors in between degree 22.5 and 292.5 were labelled defecting. By means of this procedure, 126 of the 133 were classified, 58 subjects as cooperative and 68 subjects as defecting.

The N-person game was presented as a decision-making experiment in which certain monetary gains were the consequence of certain choices. The instructions were given to the subjects by an audio-tape recorder and the subjects could simultaneously read a written version. The instructions were carefully worded to avoid suggesting any choice principle. During the instruction and the decision-making trials, each subject had available a copy of the payoff matrix of the game as presented in Table 1. The payoffs represent Dutch quarters ( $f0.25$ ).

Subjects were told that at the beginning of the experiment each subject had a credit of 60 Dutch quarters and that they could win or lose quarters depending on their own decisions and on the decisions of the other subjects. They were told that, as a payment of participation in the experiment each subject would receive the 60 quarters plus (or minus) the gains (or losses) in the experiment. It was carefully explained how the subjects could look up the monetary consequences of certain decisions in the payoff matrix. The subjects were informed that the experiment consisted of eight decision-making trials. They were told that only after the third trial would they receive information on collective choice behavior in the first, second, and third trial. It was stressed that all choices would be kept strictly confidential: no one but the subject himself and the experimenter would ever know the subject's choice behavior and his payment for the experiment.

After the explanation of the game a quiz was administered to ensure complete understanding of the task. All the incorrect answers were explained by the experimenter. Before the first trial, the subjects had to report in writing how many times they expected each of the other subjects to choose C or D in the coming eight trials. This constituted the measure of a subject's expectations about the choice behavior of others before feedback. Next, the first trial began. Each player filled out his choice on a choice form that was collected by the experimenter after every trial. The proportion of cooperative strategy choices over the first three trials was the dependent variable.

After the choices for the third trial were collected, the experimenter "counted" the choices that were made during the first, second, and third trial and gave bogus feedback. Groups were randomly assigned to one of the two feedback conditions. Each group in the majority cooperative

condition (i.e., majority C) was told that in the first and in the third trial, six subjects had chosen C and two subjects had chosen D; whereas in the second trial, seven subjects had chosen C and one subject had chosen D. Analogous feedback was given in the majority defecting condition (i.e., majority D), except that this time the C - D choice pattern was reversed.

After the feedback induction, the subjects had to report in writing how many times they expected each of the other subjects to choose C in the remaining five trials. This constituted the measure for subjects' expectations about the choice behavior of others after feedback. Finally, the choices for the fourth to the eighth trial were collected. The proportion of cooperative strategy choices over these five trials was used as the main measure of choice behavior after feedback.

At the end of the experiment, it was checked whether the subjects correctly understood the structure of the game and whether the feedback manipulation had been successful.

As a check on the correct understanding of the game, the subjects were asked what strategy would yield the highest payoff (1) in case a majority of the players chooses C; and (2) in case a majority of the players chooses D. The first question was answered correctly by 97% of the 126 subjects, and the second question was answered correctly by 96%.

After the first three trials subjects received bogus feedback about the choices group members had made in the previous trials. For each of the preceding trials, subjects were asked whether a majority of the players chose C, D, or as much C as D. The results on these questions indicated an almost complete remembering of the choices made by the group, that is, 100%, 98%, and 99% of the subjects recalled correctly the announced decisions in the respective trials.

The above results suggest that subjects correctly understood the structure of the game and that the feedback manipulation was successful.

## RESULTS

The design of our study is a partially hierarchical design in which subjects were randomly assigned to decision-making groups in which no interaction between subjects was allowed. In this design, the groups constitute a random factor nested under the factors feedback condition and type of game, whereas all factors were crossed with value orientation. Preliminary tests of (1) the effect of differences between the

decision-making groups, (2) the group by value orientation interaction, and (3) the three-way value orientation by feedback by type of game interaction, yielded insignificant effects at an alpha level of .10 on subjects' decisional behavior both before and after feedback. Although they were included in the design, in the following the insignificance of these three effects will not be reported again.

### CHOICE BEHAVIOR BEFORE FEEDBACK

First, we carried out an analysis of variance based on the arcsine transformation<sup>1</sup> of the proportions of D-choices in the first three trials, testing the effect of type of game (NPD; NCD; NTD) and value orientation (cooperators; defectors).

Hypothesis 1 received support. This appeared from a significant effect for type of game [ $F(2,90) = 24.47, p < .001$ ] and from a Dunn multiple comparison test that showed on a .05 level of significance that the proportion defecting choices in NPD ( $M = .54$ ) is greater than the proportion in NCD ( $M = .33$ ), whereas the latter is greater than the proportion in NTD ( $M = .06$ ).

Also hypothesis 2 was confirmed. There was a significant effect for value orientation [ $F(1,90) = 9.42, p < .005$ ] indicating (see Table 2) that defectors chose more D ( $M = .40$ ) than cooperators ( $M = .21$ ). The value orientation by type of game interaction failed to reach the significance level [ $F(2,90) = 3.00, p = .055$ ]. With the help of the Dunn multiple comparison test (Dunn, 1961) it appeared that, for the .05 level of significance, defectors chose more D in NPD ( $M = .68$ ) than in NCD ( $M = .41$ ) and NTD ( $M = .08$ ). Cooperators on the other hand chose more D in NPD and in NCD ( $M = .31$  and  $M = .27$ , respectively) than in NTD ( $M = .04$ ). These results are discussed in more detail in the discussion section.

Subjects' expectations are shown in Table 3. An analysis of variance with subjects' arcsine transformed expectations before the first trial as dependent variable yielded a significant effect for value orientation [ $F(1,90) = 5.29, p < .025$ ] and for type of game [ $F(2,90) = 19.34, p < .001$ ]. The interaction effect was not significant. Defectors expected more D-choices than cooperators ( $M = .38$  versus  $M = .29$ ). Furthermore, the proportion of expected D-choices in NPD ( $M = .48$ ) is higher than the proportion in NCD ( $M = .35$ ), which, in its turn, is higher than the proportion in NTD ( $M = .18$ ; Dunn multiple comparison test,  $\alpha = .05$ ).

1. Analyses of variance on the original (not transformed) dependent variables yielded exactly the same significant effects both before and after feedback.

TABLE 2  
Average Proportion of Defecting Strategy Choices  
Before Feedback

<i>Value Orientation</i>	<i>Type of Game</i>			<i>Mean</i>
	<i>NPD</i>	<i>NCD</i>	<i>NTD</i>	
Cooperators	.31	.27	.04	.21
Defectors	.68	.41	.08	.40
Mean	.54	.33	.06	.32

The correlations between the proportion of defecting choices actually made and the proportion of defecting choices expected of others are strongly positive; for NPD, NCD, and NTD, they are, respectively, .81, .48, and .65. For defectors this correlation was .80 and for cooperators this correlation was .65.

#### CHOICE BEHAVIOR AFTER FEEDBACK

Next, subjects' choices made in trials 4 to 8 were analyzed using the same procedures as was done with subjects' choices before feedback.

Hypothesis 1 was supported again. This appeared from a significant game effect [ $F(2,90) = 29.27, p < .001$ ]. Again NPD ( $M = .58$ ) evoked a higher proportion of D-choices than NCD ( $M = .30$ ), whereas NCD elicited more D-choices than NTD ( $M = .11$ ).

Also hypothesis 2 received support. The value orientation effect was significant [ $F(1,90) = 20.14, p < .001$ ], suggesting that after feedback defectors also made more D-choices than cooperators (see Table 4).

Hypothesis 3—the feedback effect—was also supported [ $F(1,90) = 9.04, p < .005$ ]. After one had heard that a majority had chosen D, one was more inclined to make a D-choice than in case one got to know that the majority had chosen C.

Of all possible interactions both the feedback by value orientation interaction [ $F(1,90) = 6.13, p < .025$ ] and the feedback by games interaction [ $F(2,90) = 3.20, p < .05$ ] were significant.

The feedback by value orientation interaction suggests that, across games, cooperators were hardly affected by feedback about what the majority had chosen ( $M = .21$  versus  $M = .19$ ), whereas defectors who had received majority D feedback were more inclined to make a D-choice than defectors who had received majority C feedback ( $M = .60$  versus  $M = .32$ , respectively).

TABLE 3  
Average Proportion of Expected Defecting Strategy Choices  
Before Feedback

<i>Value Orientation</i>	<i>Type of Game</i>			<i>Mean</i>
	<i>NPD</i>	<i>NCD</i>	<i>NTD</i>	
Cooperators	.38	.33	.15	.29
Defectors	.54	.38	.21	.38
Mean	.48	.35	.18	.34

TABLE 4  
Average Proportion of Defecting Strategy Choices  
After Feedback

<i>Value Orientation</i>	<i>Feedback</i>	<i>Type of Game</i>			<i>Mean</i>
		<i>NPD</i>	<i>NCD</i>	<i>NTD</i>	
Cooperators	Majority 'D'	.47	.13	.06	.21
	Majority 'C'	.26	.30	.04	.19
Defectors	Majority 'D'	.89	.49	.26	.60
	Majority 'C'	.52	.37	.08	.32
Mean for Feedback	Majority 'D'	.72	.25	.18	.40
	Majority 'C'	.42	.34	.06	.27
Mean for Type of Game		.58	.30	.11	.33

The feedback by games interaction is primarily due to the difference between the feedback conditions for the NPD game ( $M = .72$  versus  $M = .42$ ; Table 4). Both two-way interactions involving feedback are discussed in further detail below.

Subjects' expectations after one had received feedback were analyzed subsequently. It appeared that in NPD more D-choices were expected than in NCD, whereas in NCD more D-choices were expected than in NTD. This can be deduced from the significant game effect [ $F(2,90) = 26.23, p < .001$ ], as well as from a Dunn multiple comparison test ( $\alpha$ -level .05). Further defectors expected more D-choices than cooperators, that is, a main effect for value orientation was obtained [ $F(1,90) = 12.80, p < .001$ ]. Majority D feedback led to a higher expectation of D-choices than majority C feedback [ $F(1,90) = 32.03, p < .001$ ].

All two-way interaction effects were significant: the value orientation by feedback interaction [ $F(1,90) = 5.24, p < .025$ ], the game by feedback interaction [ $F(2,90) = 7.63, p < .001$ ], and the value orientation by game interaction [ $F(2,90) = 3.52, p < .05$ ].

TABLE 5  
Average Proportion of Expected Defecting Strategy Choices  
After Feedback

<i>Value Orientation</i>	<i>Feedback</i>	<i>Type of Game</i>			<i>Mean</i>
		<i>NPD</i>	<i>NCD</i>	<i>NTD</i>	
Cooperators	Majority 'D'	.57	.29	.22	.36
	Majority 'C'	.37	.35	.05	.24
Defectors	Majority 'D'	.82	.32	.55	.62
	Majority 'C'	.45	.31	.17	.31
Mean for Feedback	Majority 'D'	.72	.30	.41	.48
	Majority 'C'	.42	.33	.12	.28
Mean for Type of Game		.58	.31	.24	.38

The value orientation by feedback interaction effect is in line with the corresponding interaction effect for actual choices. Cooperators appeared to be less sensitive to the feedback inducement than defectors.

The significant game by feedback interaction may be explained by the fact that in NPD and NTD, majority D compared with majority C resulted in a higher expected proportion of D-choices [ $M = .72$  versus  $M = .42$  (NPD) and  $M = .41$  versus  $M = .12$  (NTD)], whereas no difference was observed in NCD ( $M = .30$  versus  $M = .33$ ). As the value orientation by game interaction suggests, this interpretation holds especially for defectors in NPD and in NTD.

In addition to the analyses presented thus far, the present design allows us to make within subjects comparisons. The advantage of this additional analysis is that it incorporates initial differences in choice behavior before feedback. For each subject we computed two difference scores, one for subjects' choices and one for subjects' expectations consisting of the average proportion in trial 4-8, minus the average proportion in trial 1-3. Those difference scores are shown in Table 6.

Two separate ANOVAs were carried out, one for each difference score. In case the difference score for subjects' choices served as dependent variable, only the main effect for feedback was significant [ $F(1,90) = 7.49$ ,  $p < .01$ ]. The feedback manipulation yielded also a significant effect for subjects' expectations [ $F(1,90) = 19.11$ ,  $p < .001$ ]. Thus, after majority D feedback, the proportion of D-choices increased ( $M = .07$ ) as well as the expected proportion of D-choices ( $M = .12$ ). After majority C feedback, both the proportion of D-choices ( $M = -.04$ ) and the expected proportion ( $M = -.05$ ) decreased. Furthermore, the ANOVA on the difference score for subjects' expectations yielded a



TABLE 6  
Average Increase (+) or Decrease (-)  
in Proportion of Defecting Strategy Choices  
After Feedback

Value Orientation	Feedback	Type of Game			Mean
		NPD	NCD	NTD	
Cooperators	Majority 'D'	.06 ( .17) <sup>a</sup>	-.04 (-.04)	.06 ( .06)	-.01 ( .05)
	Majority 'C'	.07 ( .01)	-.10 ( .02)	-.03 (-.08)	-.03 (-.02)
Defectors	Majority 'D'	.17 ( .24)	.01 ( .02)	.19 ( .32)	.14 ( .21)
	Majority 'C'	-.12 (-.04)	-.01 (-.11)	.00 (-.03)	-.04 (-.06)
Mean for Feedback	Majority 'D'	.13 ( .21)	-.03 (-.02)	.13 ( .21)	.07 ( .12)
	Majority 'C'	-.05 (-.02)	-.05 (-.06)	-.01 (-.06)	-.04 (-.05)
Mean for Type of Game		.04 ( .10)	-.04 (-.04)	.04 ( .05)	.02 ( .04)

a. Numbers in parentheses are corresponding indexes for subjects' expectations.

significant effect for game [ $F(2,90) = 4.47, p < .05$ ], a significant feedback by game interaction [ $F(2,90) = 3.23, p < .05$ ], and a significant feedback by value orientation interaction [ $F(1,90) = 4.67, p < .05$ ]. These findings support our previous conclusion that defectors are more sensitive than cooperators to majority D feedback, especially in NPD and NTD. Furthermore, this analysis shows that the remarkable decrease in proportion of D-choices for cooperators in NCD after majority D feedback [ $M = .13$  (Table 4) versus  $M = .27$  (Table 2)], can be attributed mainly to initial differences in choice behavior between feedback conditions.

Finally, we estimated the relative size of the weight  $w_1$  for fear ( $r_3$ ) and the weight  $w_2$  for greed ( $r_4$ ) in Harris's K-index:  $1 - w_1(r_3) - w_2(r_4)$ , under the constraint that  $ee'$  is minimal in the model:

$$Rw = y + e$$

in which matrix  $R$  consists of  $r_3$  and  $r_4$  for each game; vector  $w$  consists of the weights  $w_1$  and  $w_2$ ; and vector  $y$  consists of the average proportion of cooperative choices before feedback (i.e., .46, .67, .94; see Table 2). Solving  $w$  in  $w = (R'R)^{-1}R'y$  yields  $w_1 = .315$  and  $w_2 = .358$ . Thus, using a quadratic loss function, in the present study it appears that the weight for greed is only slightly higher than the weight for fear.

## DISCUSSION

Our hypotheses were all confirmed. In order to predict the proportion of cooperative choices in the three games, we used separate estimates of the motivational forces fear and greed (Harris, 1969; Komorita et al., 1980). As predicted, we did indeed observe that subjects made D-choices more often in an NPD than in an NCD, whereas in the latter more D-choices were observed than in the NTD. Support for this hypothesis appears both from the analysis of actual choice behavior before, as well as after feedback. However, contrary to the findings of Alcock and Mansell (1977) and Komorita et al. (1980), we did not find that the weight for fear in  $K = 1 - w_1(\text{fear}) - w_2(\text{greed})$  has to be smaller than the weight for greed. In the present research it seems that fear and greed were equally important. More conclusive data on the question whether fear or greed predominantly causes defection can be obtained by using a special NCD and NTD game format. For the NCD one can use an almost horizontal payoff function for C, whereas the payoff function for D lies above the C-function, except in the case where no cooperative choices are made. In such a format fear would be absent. The reverse pattern of payoff functions can be constructed for NTD, yielding a game in which greed would be absent.

Consistent with Hypothesis 2, we observed that defectors made more D-choices than cooperators. This appeared from the results of the choice behavior before and also after feedback. As such, these results imply a further external validation of value orientation as a relevant factor in game behavior. The absence of a difference between defectors ( $M = .04$ ) and cooperators ( $M = .08$ ) in NTD before feedback may possibly be attributed to the highly attractive C-choice in NTD. From these results it may be concluded that except in games in which there is a relatively low conflict between the collective and the private interest (like in the present operationalization of NTD), defectors are more inclined to make D-choices than cooperators.

The group does have an effect on individual choice behavior. Majority D feedback led to more D-choices than majority C feedback. As such, this result does not only support hypothesis 3, but it also contributes to our insight of how social dilemmas may be resolved. The knowledge that a majority either does or does not contribute to a collective good seems to be sufficient to elicit a tendency to contribute. However, this is a conditional conclusion. It appeared that the cooperators' choices were not affected, whereas only defectors increased their number of D-choices after majority D feedback. Apparently, being

informed that the majority made D-choices facilitated the defectors' natural inclination to prefer a D-choice over a C-choice. Or in other words, hearing that others preferred a D-choice may have decreased a socially induced inhibition to make a D-choice.

The subjects' expectations about the choice behavior of others, a measure that was taken before trial 1 and trial 4, were highly in agreement with subjects' actual choices. This observation is sustained by several results. Both for actual choice as well as for expected choice by others, the analyses of the data before feedback indicate a similarly interpretable main effect for type of game and for value orientation. This suggests that when one expected a D-choice by others, one also made a D-choice. Therefore, it is no surprise that all relevant correlations were highly positive.

A similar pattern of results was observed after feedback. Moreover, after feedback we consistently observed a value orientation by feedback interaction, and a game by feedback interaction, indicating that cooperators were less sensitive to majority feedback, while in the majority D condition, especially in NPD and NTD, defectors increased the number of D-choices as well as their expectations that others would also make a D-choice. This was not the case in the majority C condition.

Of relevance to the present research is Kelley and Stahelski's (1970) triangle hypothesis. Basically, that hypothesis states that cooperative persons believe that others are heterogeneous with respect to their cooperativeness versus competitiveness, whereas competitive persons (defectors) believe that others are homogeneously competitive (Kelley and Stahelski, 1970: p. 77). A second implication of their hypothesis is that more than competitors, cooperative players exhibit behavioral assimilation to the behavior of others. However, in the present research we did not observe the hypothetical pattern of data concerning expected choices of others. In fact, Kelley and Stahelski's triangularity index, computed for the data in Table 3, is negative for each type of game. This finding indicates that if there were any triangular pattern, it is opposite to the predicted one. Secondly, as the data in Table 6 illustrate we did not observe that cooperators more than defectors assimilate to the defecting strategy of others. Hence, the present findings, as well as those reported in several other studies (Miller and Homes, 1975; Kuhlman and Wimberley, 1976; Messé and Sivacek, 1979), indicate that the validity of the triangle hypothesis may be restricted to the traditional 2-person, 2-alternative PDG.

In general a strong correspondence between own choices and the expected choices of others was observed. However, the question

remains whether actual choice is based on expectations of others' choices or the reverse. It seems most likely that subjects favored a reciprocity norm: when others are expected to be more cooperative, one is also more inclined to be cooperative, and the reverse. Referring to ideas of Hobbes and Hume, Taylor (1976) summarizes this point of view rather precisely when he states that an individual "should cooperate if others do, but otherwise he should not cooperate," a tendency he defines as "conditional cooperation." This interpretation comes close to Brewer's (1981) notion about trust in reciprocity. That is, other members of the group reciprocate if one exercises trusting or cooperative behavior. Brewer referred to this belief as "depersonalized trust" in the sense that it is an assumption made in the absence of knowledge of the behavior of others in the group. Gamelike properties may increase or decrease the tendency to trust other group members. Clearly, cooperators are more inclined to show depersonalized trust than defectors, and feedback about which choices are made by the majority of the group may increase or decrease the tendency to cooperative behavior.

## REFERENCES

- ALCOCK, J. E. and D. MANSELL (1977) "Predisposition and behavior in a collective dilemma." *J. of Conflict Resolution* 21: 443-457.
- ALLEN, V. L. (1965) "Situational factors in conformity," in L. Berkowitz (ed.) *Advances in Social Psychology*. New York: Academic Press.
- BEM, D. J. and C. G. LORD (1979) "Template matching: a proposal for probing the ecological validity of experimental settings in social psychology." *J. of Personality and Social Psychology* 37: 833-846.
- BONACICH, P., G. H. SHURE, J. P. KAHAN and R. J. MEEKER (1976) "Cooperation and group size in the n-person prisoner's dilemma." *J. of Conflict Resolution* 20: 687-705.
- BREWER, M. B. (1981) "Ethnocentrism and its role in interpersonal trust," in M. B. Brewer and B. E. Collins (eds.) *Scientific Inquiry and the Social Sciences*. San Francisco: Jossey-Bass.
- CALDWELL, M. D. (1976) "Communication and sex effects in a five-person prisoner's dilemma." *J. of Personality and Social Psychology* 33: 273-281.
- COLMAN, A. M. (1982) "Game theory and experimental games: the study of strategic interaction." *International Series in Social Psychology*, volume 4. Oxford: Pergamon Press.
- DAWES, R. M. (1975) "Formal models of dilemmas in social decision making," in M. F. Kaplan and S. Schwartz (eds.) *Human Judgment and Decision Making*. New York: Academic Press.
- (1980) "Social dilemmas." *Annual Review of Psychology* 31: 169-193.

- , J. DELAY, and W. CHAPLIN (1974) "The decision to pollute." *Environment and Planning* 6: 3-10.
- DAWES, R. M., J. McTAVISH, and H. SHAKLEE (1977) "Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation." *J. of Personality and Social Psychology* 35: 1-11.
- DAWES, R. M. and J. ORBELL (1981) "Social Dilemmas," in G. M. Stephenson and J. H. Davis (eds.) *Progress in applied social psychology*, volume 1. Chichester: Wiley.
- DUNN, O. J. (1961) "Multiple comparisons among means." *J. of the Amer. Stat. Assn.* 56: 52-64.
- FESTINGER, L. (1950) "Informal social communication." *Psych. Rev.* 57: 271-282.
- (1954) "A theory of social comparison." *Human Relations* 7: 117-140.
- GRIESINGER, D. W. and J. W. LIVINGSTON, Jr. (1973) "Toward a model of interpersonal motivation in experimental games." *Behavioral Science* 18: 173-188.
- HAMBURGER, H. (1979) *Games as Models of Social Phenomena*. San Francisco: W. H. Freeman and Company.
- HARDIN, G. and J. BADEN (1977) *Managing the Commons*. San Francisco: Freeman.
- HARRIS, R. J. (1969) "A geometric classification system for  $2 \times 2$  interval-symmetric games." *Behavioral Science* 14: 138-146.
- (1972) "An interval scale classification system for all  $2 \times 2$  games." *Behavioral Science* 17: 371-383.
- KAHAN, J. P. (1974) "Rationality, the Prisoner's Dilemma and population." *J. of Social Issues* 30: 189-209.
- KELLEY, H. H. and J. GRZELAK (1972) "Conflict between individual and common interest in an n-person relationship." *J. of Personality and Social Psychology* 21: 190-197.
- KELLEY, H. H. and A. J. STAHELSKI (1970) "Social interaction basis of cooperators' and competitors' beliefs about others." *J. of Personality and Social Psychology* 16: 66-91.
- KIESLER, C. A. (1969) "Group pressure and conformity," in J. Mills (ed.) *Experimental Social Psychology*. New York: Macmillan.
- KLANDERMANS, P. G. (1983) *Participatie in en sociale beweging*. Amsterdam: V.U. Boekhandel/Uitgeverij.
- KOMORITA, S. S. (1976) "A model of the n-person dilemma-type game." *J. of Experimental Social Psychology* 12: 357-373.
- , J. SWEENEY, and D. A. KRAVITZ (1980) "Cooperative choice in the N-person dilemma situation." *J. of Personality and Social Psychology* 38: 504-516.
- KUHLMAN, D. M. and A. MARSHELLO (1975) "Individual differences in game motivation as moderators of preprogrammed strategic effects in prisoner's dilemma." *J. of Personality and Social Psychology* 32: 922-931.
- KUHLMAN, D. M. and D. C. WIMBERLEY (1976) "Expectations of choice behavior held by cooperators, competitors, and individualists across four classes of experimental game." *J. of Personality and Social Psychology* 34: 69-81.
- LIEBRAND, W.B.G. (1983) "A classification of social dilemma games." *Simulation and Games* 14: 123-138.
- (1984) "The effect of social motives, communication and group size on behavior in an n-person multi-stage mixed-motive game." *European J. of Social Psychology* 14: 239-264.
- and G. J. VAN RUN (1985) "The effect of social motives across two cultures on behavior in social dilemmas." *J. of Experimental Social Psychology*.

- MAKI, J. E. and C. G. McCLINTOCK (1983) "The accuracy of social value prediction: actor and observer influences." *J. of Personality and Social Psychology* 45, 4: 829-838.
- McCLINTOCK, C. G. (1972) "Social motivation—A set of propositions." *Behavioral Science* 17: 438-454.
- , D. M. MESSICK, D. M. KUHLMAN, and F. T. CAMPOS (1974) "Motivational basis of choice in three choice decomposed games." *J. of Experimental Social Psychology* 9: 572-590.
- MESSÉ, L. A. and J. M. SIVACEK (1979) "Predictions of other's responses in a mixed-motive game: Self-justification or false consensus?" *J. of Personality and Social Psychology* 37: 602-607.
- MESSICK, D. M. and M. B. BREWER (1983) "Solving social dilemmas: a review," in L. Wheeler (ed.) *Review of Personality and Social Psychology*. Beverly Hills, CA: Sage.
- MESSICK, D. M. and C. G. McCLINTOCK (1968) "Motivational basis of choice in experimental games." *J. of Experimental Social Psychology* 4: 1-25.
- MESSICK, D. M., H. WILKE, M. BREWER, R. KRAMER, P. ENGLISH-ZEMKE, and L. LUI (1983) "Individual adaptations and structural change as solutions to social dilemmas." *J. of Personality and Social Psychology* 44: 294-309.
- MEUX, E. P. (1973) "Concern for the common good in an n-person game." *J. of Personality and Social Psychology* 28: 414-418.
- MILLER, D. T. and J. G. HOLMES (1975) "The role of situational restrictiveness on self-fulfilling prophecies: a theoretical and empirical extension of Kelley and Stahelski's triangle hypothesis." *J. of Personality and Social Psychology* 31: 661-673.
- PRUITT, D. G. (1967) "Reward structure and cooperation: The decomposed prisoner's dilemma game." *J. of Personality and Social Psychology* 7: 21-27.
- RAPOPORT, A. (1967) "A note on the index of cooperation for Prisoners' Dilemma." *J. of Conflict Resolution* 11: 101-103.
- and M. GUYER (1966) "A taxonomy of  $2 \times 2$  games." *General Systems* 11: 203-214.
- RUTTE, C. G. and H. A. M. WILKE (1984) "Social dilemmas and leadership." *European J. of Social Psychology*.
- SCHROEDER, D. A., T. D. JENSEN, A. J. REED, D. K. SULLIVAN, and M. SCHWAB (1983) "The actions of others as determinants of behavior in social trap situations." *J. of Experimental Social Psychology* 19: 522-539.
- SHAW, E. M. (1976) *Group Dynamics: The Psychology of Small Group Behavior*. New York: McGraw-Hill.
- TAYLOR, M. (1976) *Anarchy and Cooperation*. London: Wiley.
- TYSZKA, T. and J. L. GRZELAK (1976) "Criteria of choice in non-constant zero-sum games." *J. of Conflict Resolution* 20: 357-376.
- van de KRAGT, A., J. M. ORBELL, and R. M. DAWES (1983) "The minimal contributing set as a solution to public goods problems." *Amer. Pol. Sci. Rev.* 77: 112-122.
- WILKE, H. A. M., W. B. G. LIEBRAND, and D. M. MESSICK (1983) "Sociale dilemmas: een overzicht." *Nederlands Tijdschrift voor de Psychologie*, 38: 463-480.

**SAGE PUBLICATIONS, INC.**  
275 South Beverly Drive  
Beverly Hills, California 90212



**SAGE PUBLICATIONS LTD**  
28 Banner Street  
London EC1Y 8QE, England