

SUMMARY

Cooperative or altruistic behavior in the absence of egoistic incentives is an issue that has puzzled many social scientists. In this book an attempt is made to gain more insight into such behavior for a specific type of situation: the social dilemma. Social dilemmas are defined here as situations in which: a.) each player within the game has an option that guarantees him a higher payoff than all other available options; and b.) there is a configuration of choices that offers every player a higher payoff than the payoff resulting when all choose their dominant option. When the game is played only once, and behavior remains anonymous, there appear to be no egoistic incentives for refraining from choosing the dominant option. However, previous experimental research has shown that even under these conditions a considerable degree of cooperative behavior may be found.

In **chapter one**, two generally suggested causes of non-egoistic behavior are introduced: altruism and norms. Altruism refers to (non-instrumental) concern for the well-being of others. Norms refer to feelings/cognitions about how one ought to behave in a specific situation. Two specifications are made. First, altruism is defined as group-altruism. Consequently, also from the altruistic point of view the own payoff is of interest: besides being an individual, one is also a group member. Second, the norms that are considered are social justice norms. This type of norm is not directed at promoting group welfare, but may do so as a side-effect.

Chapter two starts by asking which of the two suggested causes, altruism or norms, offers a better point of departure for a theory of non-egoistic behavior in social dilemmas. In **studies 1 and 2**, valuing adherence to an equal-sharing norm and concern for the collective payoff are compared as possible motives in addition to own gain maximization. The results of the two studies are (more) supportive of the normative than of the altruistic explanation. However, since the theorizing was still very informal, this conclusion is open to criticism. That is why in **chapter three** both the altruistic and the normative line of explanation are formalized. The altruistic model draws heavily on the work of Margolis (1982). Two central features of Margolis' model are: a.) the utility of an alternative for the group and for oneself are not combined additively, but in an interactive function: the (weighted) ratio of the two marginal utilities is assumed to be important; b.) the weight that an individual assigns to self-interest in this ratio is assumed to be a

positive function of the portion of his budget that he has already spent altruistically.

The basic assumption of the normative model, which we named the Constrained Egoism Model, is that people value fair own behavior. This means that exploiting the other players in the social dilemma is accompanied by (immaterial) costs. It is assumed that the disutility of violating other players entitlements is proportional to the degree of violation. However, giving others more than they are entitled to is assumed to have no additional utility from a social justice point of view. Like Margolis' model, the Constrained Egoism Model also assumes that own payoffs are important. The two utility arguments are combined in an additive function.

In **study 3**, two sets of conflicting hypotheses are tested. Whereas Margolis' model predicts that, all else being equal, people will behave more altruistically to the degree that such behavior has more positive consequences for the group, the Constrained Egoism Model predicts that the reverse effect will occur. The larger the benefits that one's own altruistic behavior bestows on others, the lower the level of altruistic behavior that is needed to reciprocate the benefits that one received from them. Obviously, the Constrained Egoism Model also predicts that, all else being equal, people will behave more altruistically to the degree that others' altruistic behavior has more positive consequences for oneself. Margolis' model leads to the prediction that the consequences of others' behavior for oneself will not affect own behavior.

The results of study 3 support the Constrained Egoism Model to a much larger degree than Margolis' model. The focus shifts therefore from comparing the two models to elaborating the Constrained Egoism Model. In **chapter four**, a finding that both models did not predict is used as a starting-point for this elaboration. In study 3, the overall degree of altruistic behavior was lower in asymmetric dilemmas than in their symmetric counterparts. It is suggested that more than one fairness criterion may be appropriate in a specific situation. Whereas in symmetric games different criteria lead to the same prescription for fair own behavior, in asymmetric dilemmas the prescriptions are likely to differ. This leads to **normative ambiguity**: uncertainty about which fairness criterion one should adhere to. An additional assumption is that in case of normative ambiguity, **egoistic criterion selection** will occur: people will tend to adhere to the fairness criterion that requires less altruistic behavior.

An interesting feature of the Cons suggests that one's behavior will not inv maximizing and adhering to a fairness c the different fairness criteria themselves. categorization may be performed in whi according to the option that they seem t one's profit, adhering to a self-advantag adhering to a more generous fairness cri this categorization should provide a goo categorization may be performed withou that are involved, because they all prese degree of altruistic behavior expected o dilemmas, however, such a categorizati criteria are considered admissible.

An attempt is made to specify th two types of social dilemmas. The first investment dilemma in which fairness i the second type is a consumption/resour distribution is important. The theoretica criteria will be considered admissible f the two criteria focuses on the **consequ** other on the **intention** that the behavior solution does rather well as a post-hoc asymmetry effect in study 3, in which a employed and where the asymmetry wa player's altruistic behavior for the colle

Since the set of criteria is not lin asymmetry within the type of dilemma, investigated whether the same set also dilemmas with other types of asymmetr costs of altruistic behavior was investig in the size of the players' budgets. The with one exception. Although subjects the options that the model allowed, the hypothesis was only partially supported did not opt for profit were divided abo options. A new finding is that in the as proportion of profit maximizers is high dilemmas. It is hypothesized that norm

An interesting feature of the Constrained Egoism Model is that it suggests that one's behavior will not involve a trade-off between profit maximizing and adhering to a fairness criterion, nor between adhering to the different fairness criteria themselves. This implies that a post-hoc categorization may be performed in which people are classified according to the option that they seem to have chosen: maximizing one's profit, adhering to a self-advantageous fairness criterion, or adhering to a more generous fairness criterion. If the model is correct, this categorization should provide a good fit. In symmetric dilemmas the categorization may be performed without specifying the fairness criteria that are involved, because they all prescribe matching the (average) degree of altruistic behavior expected of others. In asymmetric dilemmas, however, such a categorization depends critically on which criteria are considered admissible.

An attempt is made to specify the set of admissible criteria for two types of social dilemmas. The first type of dilemma is a production/investment dilemma in which fairness in exchange is important, while the second type is a consumption/resource dilemma in which fairness in distribution is important. The theoretical analysis suggests that two criteria will be considered admissible for each type of dilemma. One of the two criteria focuses on the **consequences** of one's behavior, the other on the **intention** that the behavior expresses. The two-criteria solution does rather well as a post-hoc explanation of the symmetry-asymmetry effect in study 3, in which an investment dilemma was employed and where the asymmetry was located in the beneficence of a player's altruistic behavior for the collective.

Since the set of criteria is not limited to a specific type of asymmetry within the type of dilemma, in **chapter five** it was investigated whether the same set also performed well in investment dilemmas with other types of asymmetry. In **study 4**, asymmetry in the costs of altruistic behavior was investigated, and in **study 5** asymmetry in the size of the players' budgets. The results are generally supportive, with one exception. Although subjects did seem to limit themselves to the options that the model allowed, the egoistic criterion selection hypothesis was only partially supported: in some cases the subjects who did not opt for profit were divided about equally between the two fair options. A new finding is that in the asymmetric dilemmas the proportion of profit maximizers is higher than in comparable symmetric dilemmas. It is hypothesized that normative ambiguity not only (often)

leads to egoistic criterion selection, but also lowers the salience of fairness considerations.

In **chapter six** the solution for the other type of dilemma, the resource dilemma, is investigated. Furthermore it is tested whether the asymmetry itself leads to a lower degree of altruistic behavior, or if this effect is indeed mediated by processes triggered by normative ambiguity: egoistic criterion selection and a lowered salience of fairness considerations. In **study 6**, asymmetry as well as symmetry in the value of resource units was investigated. The results of this study are highly supportive of the Constrained Egoism Model. The fit of the post-hoc categorization is satisfactory. Also, when the asymmetry did not result in normative ambiguity, the average degree of altruistic behavior was the same as in a symmetric dilemma in which one's resource units had the same value. Furthermore, in the conditions in which the asymmetry did result in normative ambiguity, the proportion of profit maximizers was higher and egoistic criterion selection occurred.

Chapter seven concentrates on an issue that repeatedly turned up in previous chapters, but did not receive full attention there: the difference between hurting others and not helping them. A new type of ambiguity, **application ambiguity**, is introduced. Application ambiguity implies that it is not clear **whether** fairness is an appropriate issue at all in the situation at hand, as opposed to the question **which** fairness criteria are appropriate. This type of ambiguity is assumed to be present in help-dilemmas, but not in hurt-dilemmas. Furthermore, even if people consider fairness criteria appropriate, in help-dilemmas coordinating one's behavior to that of the others (a requirement to give others their due and no more) is assumed to be more difficult, on account of the absence of a focal point for tacit coordination. Such a point is assumed to be present in hurt-dilemmas. Predictions that are based on these hypotheses are tested in **study 7**. The results are suggestive, but not conclusive.

Finally, in **chapter eight** a more extensive summary of the studies and their results is given. The relevance of the conclusions, based on all seven studies, for the emergence of non-egoistic motives is briefly evaluated. The Constrained Egoism Model is also compared with another recently proposed theory on non-egoistic behavior, that of Caporael et al. (1989).

Reeds vele sociale wetenschappers hebben het verschijnsel van coöperatief of altruïstisch gedrag strijdig lijkt te zijn met het individueel egoïsme. Dit wordt een specifieke categorie van sociale dilemma's genoemd en wel de categorie van sociale dilemma's met een dominante optie. Hier gedefinieerd als situaties waarin a. iemand beschikt over een optie die hem/haar althans evenveel oplevert als andere optie dan ook, en b. er een configuratie van belangen is waarbij iedereen een hogere opbrengst krijgt door samen te werken als allen hun dominante optie kiezen. In deze situatie zal zich éénmalig voordoen en de eigen optie er geen egoïstische redenen te bestaan om niet te kiezen voor de dominante optie. Desalniettemin werd in deze studies zelfs onder deze omstandigheden een aantal mensen voor het gedrag geconstateerd.

In **hoofdstuk 1** wordt de vraag geïdentificeerd hoe het gedrag het beste verklaard kan worden. De belangrijkste mogelijke oorzaken naar voren: altruïsme, egoïsme en sociale normen. Terwijl altruïsme refereert aan (niet-instrumenteel) handelen van anderen, refereren normen aan gevoelens van schuld die zich in een bepaalde situatie behoort te voelbaar te maken. Deze specificaties geïntroduceerd. Ten eerste wordt de theorie van groepsaltruïsme. Dit impliceert dat ook de eigen uitkomsten van belang zijn voor de keuze van een lid van de groep waar het om gaat. Het is een keuze gemaakt voor een specifiek type van gedrag dat rechtvaardigheidsnormen. Dit type normen refereert aan het gemeenschappelijke belang als een norm die men toch bevorderen.

In **hoofdstuk 2** wordt een eerste poging gemaakt om welk van de twee uitgangspunten een beter model is voor het verklaren van niet-egoïstisch gedrag in sociale dilemma's. In **studies 1 en 2** worden groepsaltruïsme en egoïsme met de "gelijk delen" norm met elkaar vergeleken. De nutsargumenten naast de eigen opbrengst van de andere partij. De normatieve verklaring in sterkere mate ondersteund. Echter, daar de theoretische onderbouwing van de informeel van aard was, valt het een en ander te betwijfelen.