Kent Academic Repository

Kerr, Norbert L. (2012) *Social Dilemmas.* In: Levine, John M., ed. Group Processes. Frontiers of Social Psychology . Taylor & Francis, New York, NY, pp. 85-110. ISBN 978-1-84872-872-1.

Downloaded from

https://kar.kent.ac.uk/41323/ The University of Kent's Academic Repository KAR

The version of record is available from

http://www.psypress.com/books/

This document version

UNSPECIFIED

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal* , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

Social Dilemmas

Norbert L. Kerr

Michigan State University & University of Kent

Date: 10/28/2010

Word count: 7886 (12029 including references)

Social Dilemmas

Intragroup cooperation, broadly conceived, is pervasive in small group contexts. The coordination of effort in task groups, the search for consensus in decision making groups, the search for mutually beneficial integrative solutions in negotiations, the give and take between leaders and followers, and many other topics considered in this volume involve choices between relatively more and less cooperative behavior. But cooperation is the focal concern in research on social dilemmas, a class of high interdependence situations which highlight the conflict between personal and collective interests. In a social dilemma, the personal rewards for competitive (usually termed *defecting*) choices are higher than for a cooperative choice, regardless of what choices others in the group make. In that narrow sense, it is personally rational to compete in social dilemmas. However, the collective and personal rewards of universal cooperation are higher than those for universal defection. So, if everyone in the group makes the "rational," defecting choice, they're all worse off than if they made the "irrational," cooperative choice. In a social dilemma, a personally rational choice is collectively irrational. These dilemmas arise frequently in social life, in problems of resource conservation (Hardin's, 1968, classic *tragedy of the commons* is a well known example), providing public goods (e.g., public radio in the US, which all, including non contributors can use), and economics (e.g., trade protectionism as defection vs. open markets as cooperation).

After several decades of research on prisoner's dilemma games (PDG; e.g., Pruitt & Kimmel, 1977), a simple 2-person social dilemma, social psychological interest in the more general N-person prisoner's dilemma (NPD), or social dilemma, began about 30 years ago, stimulated by Dawes' (1981) and Messick and Brewer's (1983) influential papers. Since then, research interest has grown steadily and rapidly (a PsycINFO search of the phrase "social dilemmas" produced 68 references in the 1980s, 162 in the 1990's, and 1044 after 2000).

The purpose of this chapter is not to comprehensively review the social dilemma literature, an unfeasible task given the vastness of the field. Fortunately, there are a number of reviews available (Agrawal, 2002; Bogaert et al., 2008; Komorita & Parks, 1996; Kollock, 1998; Kopelman et al., 2002; Kerr & Park, 2001; Ledyard, 1995; Weber et al., 2004). Rather, given this volume's title and objectives, my purpose is to describe some of the groundbreaking work going on "at the frontier" of social dilemma inquiry. This sampling of cutting edge work is admittedly selective and idiosyncratic. Moreover, we will not really know for decades which of the frontiers now being explored will actually yield the most useful scientific knowledge. So perhaps it is more accurate to say that the chapter will focus on four broad topics that strike this observer as exciting—1) fuzzy social dilemmas, 2) sanctioning systems, 3) selective play environments and partner choice, and 4) subtle determinants of dilemma perception and behavior. Some other cutting-edge topics that could have as easily been featured in the chapter (and probably would have been, with a different author) will be noted at the end of the chapter.

Some scholars (Messick & Brewer, 1983; Nemeth, 1972) have suggested that early prisoner's dilemma research lost steam when it became too focused on small experimental variations within an already narrow paradigm. One theme that characterizes all of the topics I consider here is that they all illustrate a movement away from narrower to broader conceptual, paradigmatic, and methodological approaches to the study of social dilemmas. For each topic, I will begin by providing some background on the topic, laying out some of the foundational

research upon which the newer work at the frontier is based. I will then describe some lines of ground-breaking research within the topic.

Fuzzy Social Dilemmas

Foundational Research

The prototypical prisoners/social dilemma study involves a well specified task or game with little or no uncertainty about game parameters (e.g., size of the group, options available to players, the interdependencies among players, value of outcomes). In real-world dilemmas, on the other hand, there is often much more ambiguity about both the environment and the people facing the dilemma—such dilemmas might generically be termed *fuzzy dilemmas* (Heckathorn, 1998). The effects of moving from well-specified to fuzzy dilemmas has been an object of considerable research attention.

The most thoroughly studied problem has been the effect of resource/environmental uncertainty, usually examined within a resource-dilemma paradigm where cooperation consists of taking or harvesting less than one might, and where collective overharvesting can lead to the collapse of the resource pool (e.g., if total harvests are greater than the size of the resource pool, no one gets anything). Uncertainty is typically manipulated by letting the size of the resource pool vary. For example, five persons might be allowed to harvest from a pool of 500 points (no environmental uncertainty) vs. one in which the pool could take on any value between 250-750 points (high uncertainty). In the latter case, the group members have to make their harvest decisions before learning the actual size of the pool. The usual finding is that mean harvest sizes (i.e., competitive behaviors) increase with environmental uncertainty. The best supported explanation for this effect is that optimism about likely environments tends to increase with environmental uncertainty (e.g., if the value could range from 250-750, one might assume, optimistically, that it will be somewhat greater than the expected value of 500; see several of the chapters in Suileman et al., 2004, for a through review of this uncertainty research).

More recent work on fuzzy dilemmas has examined the moderating role of environmental uncertainty and the effects of other kinds of uncertainty. Particularly interesting in the latter regard has been the work by van Lange and his colleagues on *negative noise*.

Research at the Frontier

<u>Environmental uncertainty as a moderating variable</u>. Not only does environmental uncertainty have a direct and negative effect on cooperation, it also seems to moderate the effects of a number of other variables. An early example is Wit and Wilke's (1998) finding that high social uncertainty (i.e., the range of likely levels of cooperation by others) undermined cooperation only when environmental uncertainty was high. A productive program of research by de Kwaadsteniet, van Dijk, and their colleagues has demonstrated several additional such moderating effects. The basic idea behind these studies is that even when group members cannot communicate, whenever possible they will rely on shared rules to tacitly coordinate their harvesting behavior. Under conditions of low environmental uncertainty (i.e., the size of the resource pool is fixed or has small variability), that rule is usually an equality rule (i.e., each group member will take his/her equal share of the known available pool). However, when environmental uncertainty is high, such coordination is more difficult (since it is unclear just what an equal share should be). de Kwaadsteniet et al. (2006) show, for example, that personal dispositions to cooperate (SVOs) have little effect on harvesting when environmental uncertainty is low, with most group members following the equal-harvest rule, but that SVOs show their usual effects when uncertainty is high (i.e., prosocials harvest less than proselves). If we have to justify our harvests to others, we are even more likely to follow the shared and defensible equal-division rule <u>if</u> it is possible to do so (i.e., low environmental uncertainty), but when it is difficult to do so (i.e., high environmental uncertainty), the more easily justified course seems to be to

simply harvest less (de Kwaadsteniet et al., 2007). This in turn suggests that uncooperative behavior (e.g., taking more than an equal share) can be more easily seen as counternormative under conditions of low uncertainty. Thus, the more one takes, the more angry and blaming other group members are likely to be. However, under high uncertainty, where large harvests are less easily classified as rule violations, group members are less likely to react with anger to larger harvests.

Other types of uncertainty. At least four other types of uncertainty have been examined in social dilemmas. One arises for step-level public goods, where a minimum total level of contributions (the "provision point," PP) is required to provide a public good. Here, higher uncertainty can be achieved via a larger range of possible provision points. Both when group members make their contribution decisions simultaneously (Wit & Wilke, 1998) or sequentially (Au, 2004) , increasing provision point uncertainty reduced cooperative behavior. A second is social uncertainty--the range of likely levels of cooperation by others. Here, contributions to a public good tend to be reduced when social uncertainty is high (Sabater-Grande and Georgantzis, 2002; Wit & Wilke, 1998; although see van Dijk et al., 2009, Exp. 3). A third type of uncertainty is outcome uncertainty, which has been examined via the range of possible values of the public good in a step-level game. As long as every possible value was more valuable than the provision point—i.e., the average group member never risked a loss as long as the public good was obtained—then outcome uncertainty appears to have little or no effect on cooperation (van Dijk et al., 1999; McCarter et al., 2010). However, if the outcome uncertainty is great enough that actual losses are possible, the resulting loss aversion seems to undermine willingness to contribute toward providing the public good (McCarter et al., 2010). A fourth and final type of uncertainty is uncertainty about the size of the group. Here, the effect of group-size uncertainty depends on game features—group size uncertainty seems to enhance cooperation in a common pool resource (CPR) dilemma (Au & Ngai, 2003; de Kwaadsteneit et al., 2008), but to reduce it in a step-level public good game (Au, 2004). These findings highlight the importance of not assuming that the relationships observed in one kind of social dilemma will always generalize to another—one can get very different results with different dilemmas (see Abele et al., in press).

Negative Noise. An interesting twist on interacting under conditions of uncertainty is provided by the recent program of research undertaken by van Lange and his colleagues on the effects of "noise" in social dilemma settings. In these studies, noise is a disconnect between what one party does or intends and what the other party experiences. For example, in a relationship, one party may intend to act cooperatively (e.g., get home in time for the special dinner the other has prepared), but fail to do so for reasons quite independent of his/her intent (e.g., get caught in a traffic jam). Van Lange and his colleagues rightly point out that in actual, real-world situations of interdependence, such disconnects clearly occur and may even be commonplace. In such cases, partners have to interpret whether apparently uncooperative behavior should be attributed to the other's unwillingness/disinclination to cooperate or to other non-dispositional factors (with less relevance to the ongoing interaction).

Some of the results of this work are unsurprising—e.g., negative noise undermines cooperation (relative to interactions that are not saddled with such unidirectional uncertainty; e.g., van Lange et al., 2002; Tazelaar et al., 2004). However, the negative effects of negative noise are not insurmountable. Several conditions are sufficient to neutralize or mitigate these effects. One is the opportunity to communicate (Tazelaar et al., 2004). As many a late spouse has discovered, it helps when one can explain that one's apparently uncooperative act has another, more benign explanation. Another is an empathic concern for one's partner (Rumble et al., in press). One is more likely to give one's partner the "benefit of the doubt" under conditions of uncertainty if one is more inclined to look at things from the partner's perspective. Finally, it

appears that those who are dispositionally more inclined to be cooperative in the first place (e.g., those with pro-social social value orientations; e.g., van Lange, 1999), are relatively more sensitive to the effects of noise (Brucks & van Lange, 2007). Less cooperative, pro-self individuals are not as reactive and retaliatory for occasional uncooperative acts with ambiguous causes than those who both are inclined to perform and expect more cooperative behavior. These studies have also linked the adverse effects of noise to malignant attributions about one's partner under noisy conditions—e.g., that the partner is less trustworthy (Klapwijk & van Lange, 2009), has a non-benign intent (Talezaar et al., 2004), or that the partner simply does not care about jointly shared resources (Brucks & van Lange, 2008).

Another interesting finding is that the presence/absence of negative noise may moderate the effects of other variables. It is widely accepted that a tit-for-tat strategy is a particularly effective way to encourage cooperation in other group members (e.g., Alexrod, 1984), but the evidence for this has come from the study of "noiseless" environments, i.e., situations in which one could safely assume that every apparent act of defection was intentional. However, in a more realistic, noisy environment, it seems that relatively more forgiving and generous strategies are more effective than strict reciprocity—a smidgen of generosity softens the risks of an "eye for an eye" (viz., of getting trapped in a cycle of joint defection) in noisy situations where one may inadvertently appear more uncooperative than one really is (Klapwijk & van Lange, 2009, 2010).

Sanctioning Systems

Foundational Research

In his seminal paper on public goods, Garrett Hardin (1968) despaired of groups' abilities to avoid the tragedy of the commons except via "mutual coercion, mutually agreed upon." Thus, from the beginning, one potentially important structural solution to social dilemmas has been the imposition of tangible incentives or punishments outside of the dilemma itself (e.g., mandatory taxation to provide public goods; fines for uncooperative behavior). Few early studies included any possibility for the imposition of sanctions. A few, more recent studies demonstrated that cooperation could be boosted and/or defection deterred by sanctioning systems, whether they were imposed by an external agent (e.g., Rapoport & Au, 2001; Wit & Wilke, 1990), by group members on one another (e.g., Caldwell, 1976; Eek et al., 2002; Fehr & Gächter, 2002; McCusker & Carnevale, 1995;Yamagishi, 1986, 1992), or by a group leader (e.g., van Vugt & De Cremer, 1999). More recent work at the frontier has explored a variety of questions stemming from this foundational research: Exactly why do such sanctioning systems work?; do they sometimes fail or even do harm (i.e., lead to less cooperation)?; can intangible, social sanctions be as effective as tangible ones and, if so, which are most effective and why?

Research at the Frontier

<u>Why do sanctions work?</u> The answer to this question would seem to be obvious—sanctions alter the incentive structure and people will naturally respond. But there are at least two routes through which sanctioning systems could work— (1) self-interest/greed makes cooperation relatively more personally attractive under sanctioning systems and (2) if one presumes that others are similarly affected, one is less fearful of exploitative choices by others when sanctions are in place. However there seems to be more to sanctioning than such rational recalculations of risks and benefits. For example, Fehr and his colleagues (e.g., Fehr & Fischbacher, 2004a) note that options to punish defectors in their studies are just as popular when one can never play again with the defector as when one can. That is, even when it is costly to punish and any improvement in the punished player's future behavior can be of no benefit to the punisher, many players still seize a punishment option. In fact, even "third parties" who are not playing the game themselves will—although at a somewhat reduced rate—incur costs to punish uncooperative parties, especially if the defector is exploitative (i.e., defects when his/her partner

cooperates; Fehr & Fischbacher, 2004b). Both results suggest that the motives for punishing and responding to punishment extend beyond enhancing one's own game payoffs.

Another cue to the roots of sanctioning's effectiveness is that punishing behavior may be mediated by moral emotions (e.g., anger; Gächter & Fehr, 1999; de Kwaadsteneit, et al., 2010; O'Gorman et al., 2005)—punishment is to some extent the result of anger and also may communicate such anger. Still another cue is that the degree of punishment is tied more closely to the defector's relative level of defection (relative to the level of cooperation of the punisher and/or the group as a whole) than to his or her absolute level of defection (e.g., Fehr & Fischbacher, 2004b; Masclet et al., 2003), suggesting that punishment is most likely when group members are violating group norms. Finally, it is those most inclined to view the social dilemma choice as a moral one, those with pro-social orientations (Liebrand et al., 1986), who are most sensitive to being sanctioned, particularly with sanctions that emphasize both rewards for cooperation and punishments for defection (Folmer & van Lange, 2007).

These varied threads of evidence have led several scholars (e.g., Fehr & Fischbacher, 2004a; Mulder, 2009) to conclude that when sanctioning systems work, they do so not only because of their direct impact on incentives and expectations of others' behavior, but also because they increase group members' moral concerns. In other words, sanctioning systems increase potential defectors' awareness that certain behavior is counternormative and their concerns about not violating local or general (e.g., reciprocity, equal outcomes) social norms, and with the social and intrapsychic costs that of violating these norms.

When and why do sanctioning systems fail? There appear to be several ways that sanctioning systems can and do fail. One familiar one is via overjustification effects (e.g., Deci & Ryan, 1985)—external incentives can sometimes undermine internal attributions for behavior (e.g., the intrinsic goal of behaving cooperatively) and hence the willingness to perform this behavior absent those incentives. The introduction of sanctioning systems also may suggest to group members that others lack sufficient internal motivation to cooperate (e.g., high concern for the collective welfare, a disinclination to exploit others). This in turn may undermine one's trust in the cooperativeness of others and, in the absence of the sanctions, one's own willingness to cooperate (Chen et al., 2009; Mulder et al., 2005, 2006). Whereas sanctioning systems can increase moral concerns, it is important that they are seen as retributive measures—that the sanctioned behaviors are in some sense immoral and that the sanctions exist to punish bad and reward good behavior (Mulder, 2009). All else being equal, relatively more severe sanctions are more likely to convey this message, particularly if one has some trust in the authorities that impose the sanctions and as long as the sanction is not viewed as excessive (Mulder et al., 2009). And, all else being equal, punishments are more likely to convey this message than are rewards (Mulder, 2008); reward systems are more likely to be used to promote voluntary rules, whereas punishment systems tend to be reserved for enforcing obligatory rules.

Sanctioning systems also can be seen as compensatory measures—as a means of compensating "exploited" parties for the costs incurred as a result of others' uncooperative behavior. When this happens, they can undermine defectors' moral concerns and hence backfire. For example, Gneezy & Rustichini (2000) found that imposing a fine on parents for picking up their children late from day care actually increased the incidence of late pickups. Here, the fines suggested to parents that they were compensating the day care for their lateness, and this perception undermined their sense that late pickups was "bad" behavior (i.e., violated a rule or norm; also see Tenbrunsel & Messick, 1999).

The use of social sanctioning systems. The foregoing research linking sanctioning systems to group members' concerns with the moral implications of their behavior strongly suggests that it is not just the tangible costs and rewards inherent in such systems that give them

their power. In addition, sanctioning systems can tell group members what behaviors are expected and approved of. If so, sanctioning systems that simply tell group members that others disapprove of certain behavior (e.g., defection in a social dilemma) may be sufficient to deter that behavior. This has been nicely demonstrated in several studies. For example, Masclet et al. (2003) found that giving group members an option to send "disapproval points" (without any material consequences) to others was sufficient to increase contributions to a public good, nearly as much as an option to send material punishments. Likewise, Carpenter et al. (2004) found that cooperation rates were boosted simply by giving group members an option of sending iconic unhappy faces to fellow members (for similar findings with other emotional cues of disapproval, see Kerr, 2009, and Wubben et al., 2009a, 2009b). Even subtle cues that imply that others might be observing, and hence are capable of approving or disapproving, seem sufficient to boost cooperation (e.g., Burnham & Hare, 2007; Mifune et al., 2010; Rigdon et al., 2009). For example, Kurzban (2001) found that brief eye contact between male participants prior to allocation decisions boosted cooperativeness in a social dilemma. Even more dramatically, Haley and Fessler (2005) found that, relative to a neutral screensaver, a screensaver image resembling a pair of eyes boosted contributions in an ultimatum game.

The psychological processes underlying such effects are not well understood. Some researchers (e.g., Hardy & van Vugt, 2006; Mifune et al., 2010) suggest that actual or potential evaluation arouses concerns for one's reputation in the group. Others (e.g., Haley & Fessler, 2005) suggest that rather automatic, evolved perceptual/judgmental modules are the proximal trigger of more cooperative behavior. Still others (e.g., Leary & Baumeister, 2000; Levine & Kerr, 2007) suggest that the key motive is to preserve positive relationships within one's group. This latter perspective assumes that group members are very sensitive to cues that suggest that they might be excluded and alter their behavior (e.g., conform to implied or familiar behavioral norms) to forestall any such exclusion. Consistent with this view, giving group members the possibility of excluding or ostracizing fellow group members reliably boosts cooperative behavior (Cinyabuguya et al., 2005; Kerr, 1999; Kerr et al., 2009; Masclet, 2003; Ouwerkerk et al., 2005; Maier-Rigaud et al., 2010).

Selective Play Environments and Partner Choice

Foundational Research

The prototypical early prisoner's/social dilemma study used a *forced play* paradigm where the interdependence structure, behavioral options, and game players were fixed and immutable. Besides being an obviously unrealistic model of most real social dilemmas, such a paradigm also precluded many interesting and feasible routes to solving dilemmas (see Hayashi & Yamagishi, 1998). Simulation and experimental studies of behavior using any of several *selective play* protocols have suggested that it may be easier to achieve higher and more sustainable levels of mutual cooperation than previously realized. Such studies have permitted greater variability in the range of choices and interpersonal relationships available to the players (Gallucci et al., 2004). For example:

- Some studies have offered players an *exit option*, permitting them to simply withdraw from the game (e.g., Orbell & Dawes, 1993), or to exit and play with a randomly- (e.g., Schuessler, 1989) or self-selected (e.g., Hayashi & Yamagishi, 1998) new partner. Generally speaking, higher levels of cooperation are achieved with than without such options, particularly for players who are inclined to be cooperative, trusting of others, and unwilling to remain in an exploitative relationship (see Yamagishi & Hayashi, 1996; Hiyashi & Yamagishi, 1998; and Gallucci et al., 2004, for reviews).

- A few studies have examined the effect of being able to adjust one's level of interdependence with others—i.e., to seek a relatively more or less interdependent relationship. When both players in dyadic games have this option, very high levels of mutual cooperation can be achieved (Yamagishi et al., 2005). When one player has the option and the other plays a consistent strategy, the option is sensibly used (i.e., the one player increases interdependence for highly cooperative others and decreases interdependence for highly uncooperative others; van Lange & Visser, 1999). On the other hand, when the other player follows a reciprocal (i.e., tit-for-tat) strategy, more cooperative players seek higher interdependence, but more competitive players seek reduced interdependence (van Lange & Visser, 1999).
- Still other studies have given players a less restricted set of behavioral options. A nice illustration of this type of study is Kurzban et al. (2001, Exp. 1) who found that giving players a chance to either increment or decrement posted intended levels of cooperation undermined cooperation, relative to a condition where players were only able to <u>increase</u> their posted intention If only the lowest intended contribution were posted for all to see, this "increase only" option could even prompt steadily increasing cooperation (Kurzban et al., 2001, Exp. 2), in line with a *minimum reciprocity rule* (Sugden, 1984), which prescribes that one cooperates at or slightly above the least cooperative group member.

More recent research at the frontier has 1) expanded the range and variety of selective-play options and 2) begun to explore what besides one's partner's known cooperativeness leads one (rightly or wrongly) to prefer (or avoid) that person in a social dilemma.

Research at the Frontier

<u>Selective play environments</u>. The foundational work reviewed above has tended to make few and narrow departures from a simple forced-choice paradigm and then explore empirically what difference they make in behavior. A rather different approach is that of evolutionary game theorists (e.g., Hofbauer & Sigmund, 1998; Nowak, 2006) who consider many varieties of selective play – for example, does it matter if one can withdraw from interaction?; what if interaction and/or reproduction is restricted only to some others?; what if such restrictions are based on spatial proximity?; what if one's past behavior establishes a reputation that can guide others' willingness to cooperate or affiliate in the future? This approach tends to be less concerned about predicting individual behavior under particular conditions and more concerned with the relative rewardingness (and hence, fitness) of alternative behavioral strategies in large populations that interact across long periods of time. The primary methods of inquiry are simulations and formal theory development.

My goal here is not to try to review or even summarize the burgeoning and complex literature applying evolutionary game theory to the analysis of human cooperation. Since most of the work is done by evolutionary and mathematical biologists, its details are well beyond the scope of this chapter (and this author's competence). I would, though, like to offer a couple of illustrations of how such work illuminates our understanding of behavior in social dilemmas.

We saw above that allowing group members to decline to play or to move to another partner generally increased the level of cooperation in social dilemmas. Hauert et al. (2007) went well beyond this conclusion by showing that the option to withdraw from the interdependent relationship may be vital for the success of punishment as a deterrent to defection. If participation in the social dilemma is compulsory, then under reasonable assumptions (e.g., a basic social dilemma structure, imitation of successful strategies, some possibility of mutation from one strategy to another) defectors will come to dominate the population, even if some people are willing to bear the cost of punishing such defectors. But if there is an option of not

playing the social dilemma--an option that has a moderate rewardingness--a strategy of cooperating plus punishing noncooperators will come to be the dominant strategy.

As noted earlier, if all members of a population must play a social dilemma, defection evolves and dominates. This remains true when one assumes occasional mutations, resulting in offspring that are somewhat more or somewhat less cooperative than their parents. Killingback et al. (2006) reported that this situation changes drastically if the population is divided into subsets (i.e., groups), the game is played not among everyone in the population but within these subgroups only, there is some small but real chance of dispersal (i.e., movement between groups). Under these assumptions, group members using strategies that maximize cooperation quickly evolve. Others (see Nowak et al., 2010, for an overview) have shown that many other variations on this subgrouping or clustering theme have similar effects, whether the clustering is based on physical proximity, on social networks, some common phenotype, or as in Killingback et al., on simple grouping of the population into subgroups.

Analyses like these are beginning to appear in the psychological literature (also see van Vugt, in this volume). An excellent example is Kameda et al. (in press), who note that the benefits of group collaboration usually show diminishing returns with group size. This is contrary to the usual presumption of social dilemmas, whereby the benefits of cooperation accumulate linearly with the number of cooperators. In both simulations and a lab experiment, Kameda et al. show that if just a few cooperators could produce a big benefit a stable population composed of both cooperators and defectors would evolve.Hence, they show how cooperation can evolve in groups without any special incentives (e.g., special pride in being a cooperator; a valuable reputation for being a cooperator; greater fitness for one's subgroup relative to other subgroups).

Choosing partners. One interesting selective-play variation is letting group members choose their fellow members. Clearly one would seek social dilemma partners who are trustworthy and avoid or abandon partners who are not. This might be determined through actual experience with the partner, but mistakes in partner selection could be a rather costly. It would be quite useful if there were some means of recognizing trustworthy partners prior to interaction using cues that are observable, reliable, and hard-to-fake (such as some aspect of physical appearance or an involuntary emotional expression; Frank, 1988). These speculations give rise to several related questions: Can people reliably distinguish between more and less cooperative partners based on initial encounters? What cues do they use to make such judgments? What is the actual diagnosticity of such cues?

An early study by Frank, Gilovich, and Regan (1993) demonstrated that a 30-min. "get to know you" session enabled their participants to estimate others' subsequent choices in a simple prisoner's dilemma game at a better than chance level. Qualitatively similar results were reported by Brown et al. (2003), who showed that opportunities to observe target persons in a variety of contexts (e.g., telling a children's story; describing one's likes and dislikes) was sufficient for the observers to give higher ratings on cooperation-relevant traits (e.g., helpful) to those targets that self-identified themselves as more altruistic.

These early studies suggested that potential partners might well provide cues to their future quality as partners, even when they are not actually functioning as partners—we humans seem to have some ability to recognize good interaction partners. However, they did not reveal just what those cues might be -- they could be any of several nonverbal cues, verbal content cues, or some mix of the two. Subsequent studies help narrow the range of possibilities. For example, at least some of the diagnostic cues seem to be contained in nonverbal or appearance behavior. Shelley et al. (2010; Exp. 2) found that naïve judges whose only information about a group of targets was gained through watching a silent videotape of the targets describing the events of the

prior day correctly judged that targets with prosocial orientations would use more cooperative strategies in making self/other allocations than would targets with proself orientations. [1] To narrow the range even further, at least some of the diagnostic cues appear to be available from still photographs. [2] Verplaetse et al. (2007) found that naïve judges could, based only on seeing still photographs, correctly classify both cooperators and defectors in a PDG at rates significantly above chance,. but only if the photos were taken at the moment the targets were actually deciding what to do in the PDG (not in a pre-game neutral photo or a practice round photo; cf. Brown et al., 2003, Exp. 2). Likewise, Shelley et al. (2010, Exp. 3) found no relationship between the actual cooperativeness of people (assessed via social value orientations) and judge's expected cooperativeness of those same people when only a still photo with a neutral expression was available to judges. These studies suggest that there may be little useful diagnostic information available in flat, neutral, or nonexpressive depictions of potential partners. But the story is very different when some emotional expression is possible. When judges are shown photos of people posing a smile (either with [Exp. 5] or without [Exp. 4] poses of other emotions), they can (to some degree, at least) correctly distinguish between more vs. less cooperative people (Shelley et al., 2010).

All this suggests 1) that we rely upon others' emotional expressions to assess their fitness as cooperation partners (the perceptual link) and 2) that at least some aspects of others' emotional expressiveness are reliably linked to their actual or likely level of cooperation (the behavioral link). There is growing evidence for both links, particularly when it comes to positive emotional expressions. For the perceptual link, people perceive those who smile, particularly those whose smiles are genuine (Duchene smiles), to be more trustworthy and concerned about others (Brown et al., 2003, Exp. 3; Krumhuber et al., 2007). People are also more willing to act cooperatively towards such smiling partners in games that require trust (Scharlemann et al., 2001; Krumhuber et al., 2007). Recall that the most useful cues for partner quality should be expressed involuntarily (Frank, 1988), because voluntarily controlled cues are more easily faked and could thus be misleading. In their analysis of silent video clips, Brown et al. (2003) found that judgments of a target's concern for others were reliably associated with a number of such involuntary cues (e.g., Duchene smiles, brief smiles, a "concern furrow" at the brow), but not with several voluntary ones (e.g., eyebrow flashes and raises, open smiles).

For the behavioral link (between emotional expression and cooperation), there is also growing evidence. All the involuntary facial cues that Brown et al.'s (2003) judges relied upon were also reliably associated with the targets' self-reported levels of altruistic behavior. And the frequency of Duchene smiles displayed in an interaction between two friends who had to split their experimental earnings was reliably associated with their expressed willingness to extend help to others (Mehu et al., 2007). In unpublished studies by Carnevale (1977) and Mills (1978) [cited by Shelley et al., 2010], silent videotapes of people describing emotionally charged experiences were judged as relatively more positive for more cooperative people and relatively more negative (angry and sad) for less cooperative people. All of these studies suggest it may be

---

[1] It is well established that social value orientations are reliably predictive of cooperative behavior in social dilemma and other mixed motive settings (e.g., Balliet et al.,2009).

[2] A series of studies by Yamagishi et al. (2003) indirectly bolster this argument. When shown pictures of former cooperators and defectors in prisoners dilemma games and subsequently asked to identify those shown and not shown, participants were better able to recognize former defectors than former cooperators. Interestingly, the false alarm rate was also higher for former defectors than cooperators. Clearly, there is something in even simple facial images of less cooperative others that makes them distinctive and memorable.

the general positivity of one's facial and nonverbal expressions that signals a more cooperative person (see Schug et al., 2010, for a competing interpretation).

Despite the null results reported above for neutral, still pictures, there are also growing indications that physical attractiveness might provide a useful cue for partner selection, at least for male partners. In an early study, Mulford et al. (1998) found that there was an association between how attractive one perceived a potential partner and how willing one was to play a prisoner's dilemma with that partner. These effects were not moderated by the sex of the judge or the sex of the potential partner. So people seem to think that attractive partners are cooperative partners. Ironically, the actual cooperation data suggest otherwise. Whether attractiveness is indexed by subjective judgments by independent judges (Takahashi et al., 2006) or by symmetry in body (Zaatari & Trivers, 2007) or face (Sanchez-Pages & Turiegano, 2010), more attractive/symmetric men are in fact less cooperative than their less attractive/symmetric brothers. And neither attractiveness (Takahashi et al., 2006) nor symmetry (Zaatari & Trivers, 2007) among females is reliably linked to their willingness to cooperate. This full pattern of results has been interpreted in evolutionary terms (e.g., see Zaatari & Trivers, 2007; Takahashi et al., 2006)—more genetically desirable (i.e., more attractive, symmetric) males may not be as dependent on cooperation to attract mates or to gain other resources (e.g., they can use aggression more effectively). This line of argument implicates other potential indicators of male reproductive or survival fitness, such as testosterone levels. There are some intriguing hints that early exposure to testosterone in utero and resulting observable morphological markers (e.g., 2D:4D ratio; masculinzed facial features) are linked to cooperativeness, but the findings are preliminary and inconsistent (cf. Millet & Dewitte, 2006, 2009; Sanchez-Pages & Turiegano, 2010; Pound et al., 2009).

Subtle Determinants of Dilemma Perception and Behavior: Priming/Framing Foundational Research

The traditional analysis of social dilemmas is an economic, rational-choice, expected-utility one--people can be expected to respond to the objective payoffs available in the situation, or at least to the subjective values that they attach to the possible outcomes. This approach suggests that situational features that do not materially alter the dilemma's incentives should not affect behavior choices. Nevertheless, much early social dilemma research showed that framing a dilemma with a fixed incentive structure in different ways could alter levels of cooperation. For example, framing a problem so that it looks as though a cooperative choice rewards others resulted in more cooperation than a functionally identical framing that looks as though a defecting choice punishes others (Komorita, 1987; Fehr & Gächter, 2000; cf. Kerr & Kaufman-Gilliland, 1997). A fairly large literature has examined framing functionally the same dilemma in public-good terms (where one must decide how much to give toward a shared resource) vs. resource-dilemma terms (where one must decide how much to take from a shared resource). More often than not, cooperation rates have been higher in the latter, "take" framing, but there are many exceptions and complications to such a simple summary (see DeDreu & McCusker, 1997; Tenbrunsel & Northcraft, 2010; and Weber et al., 2004, for reviews).

An appealing alternative to the traditional, rational choice perspective is Weber, Koppelman, and Messick's (2004) appropriateness framework. The latter suggests that behavioral choices are often governed by one's assessment of just what is the most appropriate way to behave in the given situation-- by how one answers the question, "what does a person like me do in a situation like this?". From the appropriateness perspective, apparently superficial aspects of the situation—for example, how it is framed or described—can determine how the situation is construed and hence what rules or norms one might follow (Tenbrunsel & Northcraft, 2010). Even the label used to describe a social dilemma can lead to very different construals and

behavior. For example, Batson and Moran (1999) found that characterizing a simple prisoner's dilemma game as a "business transaction" resulted in less cooperation than when it was described as a "social exchange". Similarly, Cronk (2007) showed that giving a trust game an extra label (this is an *osatua* game) that was associated among his participants (Maasii men) with need-based giving resulted in giving less and returning less in a trust game than occurred with without such a label. Such associations need not be longstanding -- Crock and Wasielewski (2008) found that having American undergraduates read a brief description of the Maasii culture and the osatua concept led them to exhibit the same framing effects as the Maasii tribesmen who grew up with the concept.

Such findings raise several interesting questions that have gained increasing research attention— for example, can such framing effects be primed?; if so, what are the concepts that effectively prime higher levels of cooperation?; and when and how do such primes work?

Research at the Frontier

If, as the Weber et al. model suggests, the particular construal one puts on a dilemma ("what kind of situation is this and how are people like me supposed to behave in it?") guides one's behavior, then making a particular construal cognitively accessible—e.g., by priming it—should guide how one construes the situation. This was demonstrated early on using fairly heavy-handed primes by Eliot, Hayward, and Canon (1998). Participants first read a set of news briefings that extolled either an entrepreneurial business strategy (i.e., emphasizing autonomy and individual achievement) or a cooperative business strategy (i.e., emphasizing teamwork and group achievement). They were also asked to provide examples of and arguments in favor of their primed strategy. In a second, ostensibly unrelated experiment, Eliot et al. found greater cooperation in a public goods game for those primed with the cooperative strategy. Other studies have shown that far more subtle primes produce similar effects. For example, Hertel and Fiedler (1994) found that priming the positive connotations of cooperation and the negative connotations of competition in an ostensive memory test increased cooperation in a social dilemma (also see Hertel & Fiedler, 1998; Utz et al., 2005). Subsequent work has shown similar effects for priming or activating business concepts (Kay et al., 2004), interdependence/independence (Utz, 2004b), morality (Utz et al., 2005), legality (Callan et al., 2010), and broad/high-level vs. narrow/low-level construal mind-sets (Sanna et al., 2009).

Kay and Ross (2003) provided nice evidence that such priming effects may indeed be mediated by group members' construals of the game. After reading cooperative or competitive words in a scrambled sentence task, participants were shown a generic prisoner's dilemma game and asked to rate how appropriate each of several alternative names would be for the game. Those primed with competition construed the game as one for which competitive behavior was appropriate (e.g., preferred *The Wall Street Game* as a label) and were themselves less willing to be cooperative. Those primed with cooperation construed the game as a more cooperative (e.g., preferred *The Team Game* as a label) and were also more willing to cooperate. The priming effects on behavior were also stronger after participants had already named the game (and settled on a construal) than when they had not yet done so.

Recent work has shown that such priming effects in social dilemmas depend upon aspects of the person being primed and aspects of the dilemma, and such moderation effects have shed considerable light on the underlying psychological processes. For example, priming effects are more pronounced when the dilemma itself is more ambiguous and lacks clear normative demands (Kay et al., 2004), that is, where the context does not already provide a compelling construal. Likewise, priming effects are more pronounced when one does not already have a chronically available construal. Smeesters et al. (2009a, Exp. 1) found that people with a strong, consistent social value orientation (SVO; regardless of whether it was pro-social or pro-self)

were not affected by primes of religiosity (toward more prosocial behavior) or business (toward less prosocial behavior) in a dictator game, but those with weaker, inconsistent SVOs were. Smeesters et al. (2009a, Exp. 2) also argued that priming the self-concept tends also to prime a set of construals of how one usually acts in various situations. In support of this idea, they also found less sensitivity to religious or business primes on prosocial behavior among those whose self concept had been previously primed.

Whether priming a concept will lead to cooperation vs. competition will depend upon the preexisting associations the target of the priming has. So, for example, people with pro-social orientations think that being smart or competent implies one would act cooperatively (as they themselves act), whereas people with pro-self orientations believe that anyone who is smart would act competitively (as they tend to act; van Lange & Kuhlman, 1994). Thus priming competence should tend to increase competitiveness for pro-selfs but increase competitiveness for prosocials, which was the pattern Utz, Ouwerkerk, and van Lange (2004) observed. Likewise, subtly activating the self (by circling first-person pronouns in a text) led to more cooperation among prosocials but to less cooperation among proselves (Utz, 2004a). Likewise, the cooperation-inhibiting effect of priming legal concepts was limited to those who already saw the world in competitive, zero-sum terms (Callan et al., 2010).

Most of the priming results we have been reviewing are fairly straightforward—priming competition or concepts associated with competition increases competitive behavior. But Smeesters, Wheeler, and Kay (2009b) suggested that there is a second interesting route for primes—they can color our perception of the other people with whom we may interact. So, for example, priming unkindness could both make a person construe the situation as one where unkind, proself behavior is appropriate and make the person see his/her interaction partner as more unkind. In this example, either kind of priming would be expected to reduce prosocial behavior. But, they suggest, which kind of priming is likely to occur depends upon how focused the person is on others—being focused outward, on others, would tend to engage the second, perceptual route. Smeesters et al. (2009b) present the results of a pair of experiments were nicely consistent with this theoretical argument.

Concluding Thoughts

The four topics we have been considering are only some of those we might have discussed. For example, with but a few exceptions, the first few decades of social dilemma research was conducted in western cultures, predominantly the U.S. and western Europe. Today, there is increasing interest in discovering more about how culture shapes responses to social dilemmas (e.g., Buchan et al., 2009; Yamagishi & Suzuki, 2010). In addition, the growing use of biopsycholgical and neuroscience methods to analyze behavior has also been reflected in the social dilemma area (e.g., Fehr, 2009; Hein & Singer, 2010). And all the questions that can be posed for individual behavior within social dilemmas—and a few that cannot—may also be examined when groups rather than individuals are the “players” (e.g., Bornstein, 2003; Wildschut & Insko, 2007). Very interesting work is also being done on the way in which key perceptions are formed, such as whether others are trustworthy (e.g., Kramer & Cook, 2004; Yamagishi, 2001) or a particular distribution of outcomes is fair (e.g., van Dijk et al., 2010). And good progress continues to be made in many traditional areas of social dilemma inquiry (e.g., social value orientations, Bogaert et al., 2008, van Lange & Joireman, 2010; the efficacy of cooperative action, Yu et al., 2009).

The scientific study of social dilemmas continues to expand and flourish. Most exciting, I think, is the truly multidisciplinary nature of this work. If one attends a social dilemma conference or simply starts browsing through the literature, one will encounter investigators from the natural sciences (e.g., biologists, ethologists, zoologists, ecologists), from the social sciences

(e.g., social and organizational psychologists, sociologists, economists, political scientists), and from many applied areas (e.g., environmentalists, engineers, conservationists, regulators, policy advisors). Besides this disciplinary diversity one will also see considerable methodological diversity. In social psychology, laboratory experiments continue to predominate, but one will also encounter simulations, ethnographic studies, field experiments, opinion and resource use surveys, archival analyses, observations of natural behavior, and comparisons across cultures and species. Such a diversity of conceptual and methodological approaches offers the promise and, increasingly, the payoff of converging evidence. The mysteries of human cooperation are not fully explored, but if the activity at the frontiers of our knowledge is any indication, we are well on the way to understanding the core mysteries--when and why we put common interest ahead of self interest.

References

Abele, S., Stasser, G., & Chartier, C. (in press). Conflict and coordination in the provision of public goods: A conceptual analysis of continuous and step-level games. *Personality and Social Psychology Bulletin*.

Agrawal, A. (2002). Common resources and institutional sustainability. In E. Ostrom, T. Dietz, *N*. Dolsak, P. C. Stem, S. Stonich, & E. U. Weber (Eds.), *The drama of the commons* (pp. 41-86). Washington, DC: National Academy Press.

Au, W. (2004). Criticality and environmental uncertainty in step-level public goods dilemmas. *Group Dynamics: Theory, Research, and Practice, 8*(1), 40-61.

Au, W. T., & Ngai, M. Y. (2003). Effects of group size uncertainty and protocol of play in a common pool resource dilemma. *Group Processes & Intergroup Relations, 6*(3), 265-283.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations, 12*(4), 533-547.

Batson, C. D., & Moran, T. (1999). Empathy-induced altruism in a prisoner's dilemma. *European Journal of Social Psychology, 29*(7), 909-924.

Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology, 47*, 453-480.

Bornstein, G. (2003). Intergroup conflict: Individual, group, and collective interests. *Personality and Social Psychology Review, 7*(2), 129-145.

Brown, W. M., Palameta, B., & Moore, C. (2003). Are there nonverbal cues to commitment? An exploratory study using the zero-acquaintance video presentation paradigm. *Evolutionary Psychology, 1*, 42-69.

Brucks, W. M., & Van Lange, P. A. M. (2007). When prosocials act like proselfs in a commons dilemma. *Personality and Social Psychology Bulletin, 33(5)*, 750-758.

Brucks, W. M., & Van Lange, P. A. M. (2008). No control, no drive: How noise may undermine conservation behavior in a commons dilemma. *European Journal of Social Psychology, 38(5)*, 810-822.

Buchan, N. R., Grimalda, G., Wilson, R., Brewer, M., Fatas, E., & Foddy, M. (2009). Globalization and human cooperation. *PNAS Proceedings of the National Academy of Sciences, 106*(11), 4138-4142.

Burnham, T. C., & Hare, B. (2007). Engineering human cooperation — Does involuntary neural activation increase public goods contributions? *Human Nature, 18*, 88−108.

Caldwell, M. D. (1976). Communication and sex effects in a five-person prisoner's dilemma game. *Journal of Personality and Social Psychology, 33*(3), 273-280.

Callan, M. J., Kay, A. C., Olson, J. M., Brar, N., & Whitefield, N. (2010). The effects of priming legal concepts on perceived trust and competitiveness, self-interested attitudes, and competitive behavior. *Journal of Experimental Social Psychology, 46*(2), 325-335.

Carnevale, P. (1977). *Cooperators, competitors and individualists encode nonverbal affect*. Paper read at the Eastern Psychological Association Convention, Boston.

Carpenter, J. P., Daniere, A. G., & Takahashi, L. M. (2004). Cooperation, trust, and social capital in southeast Asian urban slums. *Journal of Economic Behavior & Organization, 55*(4), 505-531.

Chen, X-P., Pillutla, M., & Yao, X. (2009). Unintended consequences of cooperation inducing and maintaining mechanisms in public goods dilemmas: Sanctions and moral appeals. *Group Processes & Intergroup Relations. 12*(2), 241-255.

Cinyabuguma, M., Page, T., Putterman, L., 2005. Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics, 89,*
1421–1435.

Cronk, L. (2007): The influence of cultural framing on play in the trust game: A Maasai example. *Evolution and Human Behavior, 28*, 352–358.

Cronk, L., & Wasielewski, H. (2008). An unfamiliar social norm rapidly produces framing effects in an economic game. *Journal of Evolutionary Psychology, 6*(4), 283-308.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169-193.

Deci, E., & Ryan, R.M., (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press, New York.

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*(5), 1093-1106.

de Kwaadsteniet, E. W., & van Dijk, E. (2010). Social status as a cue for tacit coordination. *Journal of Experimental Social Psychology, 46*(3), 515-524.

de Kwaadsteniet, E. W., van Dijk, E., Wit, A., & de Cremer, D. (2006). Social dilemmas as strong versus weak situations: Social value orientations and tacit coordination under resource size uncertainty. *Journal of Experimental Social Psychology, 42*(4), 509-516.

de Kwaadsteniet, E. W., van Dijk, E., Wit, A., De Cremer, D., & de Rooij, M. (2007). Justifying decisions in social dilemmas: Justification pressures and tacit coordination under environmental uncertainty. *Personality and Social Psychology Bulletin, 33*(12), 1648-1660.

de Kwaadsteniet, E. W., van Dijk, E., Wit, A., & De Cremer, D. (2008). 'How many of us are there?': Group size uncertainty and social value orientations in common resource dilemmas. *Group Processes & Intergroup Relations, 11*(3), 387-399.

de Kwaadsteniet, E. W., van Dijk, E., Wit, A., & De Cremer, D. (2010). Anger and retribution after collective overuse: The role of blaming and environmental uncertainty in social dilemmas. *Personality and Social Psychology Bulletin, 36*(1), 59-70.

Elliott, C. S., Hayward, D. M., & Canon, S. (1998). Institutional framing: Some experimental evidence. *Journal of Economic Behavior and Organization, 35(4),* 455—464.

Eek, D., Loukopoulos, P., Fujii, S., & Gärling, T. (2002). Spill-over effects of intermittent costs for defection in social dilemmas. *European Journal of Social Psychology, 32*(6), 801-813.

Fehr, E. (2009). *Social preferences and the brain*. San Diego, CA, US: Elsevier Academic Press.

Fehr, E. & Fischbacher, U. (2004a). Social norms and human cooperation. *TRENDS in Cognitive Science, 8*(4), 185-190.

Fehr, E. & Fischbacher, U. (2004b). Third-party punishment and social norms. *Evolution and Human Behavior, 25*, 63-87.

Fehr, E., Gächter, S. (2000). Do incentive contracts crowd out voluntary cooperation? Working Paper No. 34, Institute for Empirical Research in Economics, University of Zurich.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137-140.

Folmer, C. P. R., & van Lange, P. A. M. (2007). Why promises and threats need each other. *European Journal of Social Psychology, 37,* 1016-1031.

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York, NY, US: W W Norton & Co.

Frank, R. H., Gilovich, T., & Regan, D. T. (1993). The evolution of one-shot cooperation: An experiment. *Ethology & Sociobiology, 14*(4), 247-256.

Gächter, S., & Fehr, E. (1999). Collective action as a social exchange. *Journal of Economic Behavior and Organization, 39*, 341–369.

Gallucci, M., van Lange, P. A. M., & Ouwerkerk, J. W. (2004). Freedom of movement: A strategic analysis of social dilemmas with the option to move. In R. Suleiman, D. V. Budescu, I. Fischer & D. M. Messick (Eds.), *Contemporary psychological research on social dilemmas*. (pp. 180-208). New York, NY, US: Cambridge University Press.

Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies, 29*(1), 1-17.

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26*(3), 245-256.

Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 1243-1248.

Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin, 32*(10), 1402-1413.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., & Sigmund, K. (2007). Via freedom to coercion: The emergence of costly punishment. *Science, 316*(5833), 1905-1907.

Hayashi, N., & Yamagishi, T. (1998). Selective play: Choosing partners in an uncertain world. *Personality and Social Psychology Review, 2*(4), 276-289.

Heckathorn, D. (1998). Social Cooperation in Fuzzy Dilemmas: Technological versus Valuational Uncertainty. Paper presented at the annual meeting of the American Sociological Association, San Francisco, CA.

Hein, G., & Singer, T. (2010). Neuroscience meets social psychology: An integrative approach to human empathy and prosocial behavior. In M. Mikulincer, & P. R. Shaver (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature.* (pp. 109-125). Washington, DC, US: American Psychological Association.

Hertel, G., & Fiedler, K. (1994). Affective and cognitive influences in a social dilemma game. *European Journal of Social Psychology. Special Issue: Affect in Social Judgments and Cognition, 24*(1), 131-145.

Hertel, G., & Fiedler, K. (1998). Fair and dependent versus egoistic and free: Effects of semantic and evaluative priming on the "ring measure of social values". *European Journal of Social Psychology, 28*(1), 49-70.

Hofbauer, J. & Sigmund, K. (1998). *Evolutionary games and population dynamics.* Cambridge, UK: Cambridge University Press.

Kameda, T., Tsukasaki, T., Hastie, R., & Berg, N. (in press). Democracy under uncertainty: The “wisdom of crowds” and the free-rider problem in group decision making. *Psychological Review.*

Kay, A. C., & Ross, L. (2003). The perceptual push: The interplay of implicit cues and explicit situational construals on behavioral intentions in the prisoner's dilemma. *Journal of Experimental Social Psychology, 39*(6), 634-643.

Kay, A. C., Wheeler, S. C., Bargh, J. A., & Ross, L. (2004). Material priming: The influence of mundane physical objects on situational construal and competitive behavioral choice. *Organizational Behavior and Human Decision Processes. Special Issue: Social Comparison Process, 95*(1), 83-96.

Kerr, N. L. (1999). Anonymity and social control in social dilemmas. In M. Foddy et al. (Eds.), *Resolving social dilemmas* (pp. 103-119). Philadelphia, PA: Psychology Press.

Kerr, N. L. (2009). Affective feedback as a means of social control in social dilemmas. Paper presented at the 13th International Conference on Social Dilemmas (ICSD2009), Kyoto, Japan, August 20-24, 2009.

Kerr, N. L. & Kaufman-Gilliland, C. (1997). Rationalizing defection in social dilemmas. *Journal of Experimental Social Psychology*, **33**, 211-230.

Kerr, N. L. & Park, E. (2001). Group performance in collaborative and social dilemma tasks: Progress and prospects. In M. Hogg & S. Tindale (Eds.), *Blackwell Handbook of Social Psychology* (Vol. 4, Groups). Blackwell: Malden, MA.

Kerr, N. L., Rumble, A. C., Park, E. S., Parks, C. D., Ouwerkerk, J. W., Gallucci, M., & van Lange, P. A. M. (2009). "One bad apple spoils the whole barrel": Social exclusion as a remedy for the One Bad Apple Effect. *Journal of Experimental Social Psychology, 45(4),* 603-613.

Killingback, T., Bieri, J. & Flatt, T. (2006). Evolution in group structured populations can resolve the tragedy of the commons. *Proceedings of the Royal Society B, 273,* 1477–1481.

Klapwijk, A., & Van Lange, P. A. M. (2009). Promoting cooperation and trust in "noisy" situations: The power of generosity. *Journal of Personality and Social Psychology, 96(1),* 83-103.

Klapwijk, A., & Van Lange, P. A. M. (2010). The greater power of generosity. Unpublished manuscript, Free University of Amsterdam.

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology, 24,* 183-214.

Komorita, S. S. (1987). Cooperative choice in decomposed social dilemmas. *Personality and Social Psychology Bulletin, 13,* 53-63.

Kopelman, S.. Weber, J. M., & Messick, D. M. (2002). Factors influencing cooperation in commons dilemmas: A review of experimental psychological research. In E. Ostrom, T. Dietz, N. Dolsak, P C. Stern, S. Stonich, & E. U. Weber (Eds.), *The drama of the commons* (pp. 113-156). Washington, DC: National Academy Press.

Kramer, R. M., & Cook, K. S. (2004). *Trust and distrust in organizations: Dilemmas and approaches*. New York, NY, US: Russell Sage Foundation.

Krumhuber, E., Manstead, A. S. R., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion, 7*(4), 730-735.

Kurzban, R. (2001). The social psychophysics of cooperation: Nonverbal communication in a public goods game. *Journal of Nonverbal Behavior, 25*(4), 241-259.

Kurzban, R., McCabe, K., Smith, V. L., & Wilson, B. J. (2001). Incremental commitment and reciprocity in a real-time public goods game. *Personality and Social Psychology Bulletin, 27*(12), 1662-1673.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology, vol. 32.* (pp. 1-62). San Diego, CA, US: Academic Press.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 111-194). Princeton, NJ: Princeton University Press.

Levine, J. M, & Kerr, N. L. (2007). Inclusion and exclusion: Implications for group processes. In A. E. Kruglanski and E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed.). New York: Guilford.

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.

Maier-Rigaud, F. P., Martinsson, P., & Staffiero, G. (2010). Ostracism and the provision of a public good: Experimental evidence. *Journal of Economic Behavior and Organization, 73*(3), 387-395.

Masclet, D., (2003). Ostracism in work teams: A public goods experiment. *International Journal of Manpower, 24*(7), 867-887.

Masclet, D., Noussair, C., Tucker, S. & Villeval, M.-C. (2003). Monetary and non-monetary punishment in the VCM. *American Economic Review, 93*, 366-380.

McCarter, M. W., Rockmann, K. W., & Northcraft, G. B. (2010). Is it even worth it? The effect of loss prospects in the outcome distribution of a public goods dilemma. *Organizational Behavior and Human Decision Processes, 111*(1), 1-12.

McCusker, C., & Carnevale, P. J. (1995). Framing in resource dilemmas: Loss aversion and the moderating effects of sanctions. *Organizational Behavior and Human Decision Processes, 61*(2), 190-201.

Mehu, M., Grammer, K., & Dunbar, R. I. M. (2007). Smiles when sharing. *Evolution and Human Behavior, 28*, 415–422.

Messick, D. M., & Brewer, M. B. (1983). Solving social dilemmas: A review. In L. Wheeler & P. Shaver (Eds.), *Annual review of personality and social psychology* (Volume 3). Beverly Hills, CA: Sage.

Mifune, N., Hashimoto, H., & Yamagishi, T. (2010). Altruism toward in-group members as a reputation mechanism. *Evolution and Human Behavior, 31*(2), 109-117.

Millet K, & Dewitte S.(2006). Second to fourth digit ratio and cooperative behavior. Biological Psychology, 71, 111-115.

Millet, K., & Dewitte, S. (2009). The presence of aggression cues inverts the relation between digit ratio (2D:4D) and prosocial behaviour in a dictator game. *British Journal of Psychology, 100*(1), 151-162.

Mills, J. A. (1978). *Social motivations and encoding and decoding of nonverbal affect*. Unpublished undergraduate thesis, University of Delaware.

Motro, U. (1991). Co-operation and defection: Playing the Field and the ESS. *Journal of Theoretical Biology*, 151, 145-154.

Mulder, L. B. (2009). The two-fold influence of sanctions on moral concerns. In D. De Cremer (Ed.), *Psychological perspectives on ethical behavior and decision making.* (pp. 169-180). Charlotte, NC, US: Information Age Publishing.

Mulder, L. B. (2008). The difference between punishments and rewards in fostering moral concerns in social decision making. *Journal of Experimental Social Psychology, 44*(6), 1436-1443.

Mulder, L. B., van Dijk, E., De Cremer, D., & Wilke, H. A. M. (2006). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology, 42*(2), 147-162.

Mulder, L. B., van Dijk, E., Wilke, H. A. M., & De Cremer, D. (2005). The effect of feedback on support for a sanctioning system in a social dilemma: The difference between installing and maintaining the sanction. *Journal of Economic Psychology, 26*(3), 443-458.

Mulder, L. B., Verboon, P., & De Cremer, D. (2009). Sanctions and moral judgments: The moderating effect of sanction severity and trust in authorities. *European Journal of Social Psychology, 39*(2), 255-269.

Mulford, M., Orbell, J., Shatto, C., & Stockard, J. (1998). Physical attractiveness, opportunity, and success in everyday exchange. *American Journal of Sociology, 103*(6), 1565-1592.

Nemeth, C. (1972). A critical analysis of research utilizing the prisoner's dilemma paradigm for the study of bargaining. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 6, pp. 203-234). New York: Academic Press.

Nowak, M. A. (2006). *Evolutionary dynamics*. Cambridge, MA: Harvard University Press.

Nowak, M. A., Tarnita, C. E., & Antal, T. (2010). Evolutionary dynamics in structured populations. *Philosophical Transactions of the Royal Society B, 365,* 19–30.

O'Gorman, R., Wilson, D. S., & Miller, R. R. (2005). Altruistic punishing and helping differ in sensitivity to relatedness, friendship, and future interactions. *Evolution and Human Behavior, 26*(5), 375-387.

Orbell, J. M., & Dawes, R. M. (1993). Social welfare, cooperators' advantage, and the option of not playing the game. *American Sociological Review, 58*(6), 787-800.

Ouwerkerk, J. W., Kerr, N. L., Gallucci, M., & Van Lange, P. A. M. (2005). Avoiding the social death penalty: Ostracism and cooperation in social dilemmas. In Williams, K. D., Forgas, J. P., & von Hippel, W. (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying.* New York: The Psychology Press.

Pound, N., Penton-Voak, I. S., & Surridge, A. K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proceedings of the Royal Society B, 276,* 153-159.

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., & Au, W. T.(2001). Penalty in common pool resource dilemmas under uncertainty. *Organizational Behavior and Human Decision Processes. Vol 85*(1), 135-165.

Rigdon, M., Ishii, K., Watabe, M., & Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology, 30*(3), 358-367.

Sabater-Grande, G., & Georgantzis, N. (2002). Accounting for risk aversion in repeated prisoners' dilemma games: An experimental test, *Journal of Economic Behavior and Organization* ***48***, 37–50.

Sanchez-Pages, S., & Turiegano, E. (2010). Testosterone, facial symmetry and cooperation in the prisoners' dilemma. *Physiology & Behavior, 99*(3), 355-361.

Sanna, L. J., Chang, E. C., Parks, C. D., & Kennedy, L. A. (2009). Construing collective concerns: Increasing cooperation by broadening construals in social dilemmas. *Psychological Science, 20*(11), 1319-1321.

Scharlemann, J. P. W., Eckel, C. C., Kacelnik, A., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology, 22,* 617–640.

Schuessler, R. (1989). Exit threats and cooperation under anonymity. *Journal of Conflict Resolution, 33*(4), 728-749.

Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T., & Bonnet, K. (2010). Emotional expressivity as a signal of cooperation. Evolution and Human Behavior, 31(2), 87-94.

Shelley, G. P., Page, M., Rives, P., Yeagley, E., & Kuhlman, D. M. (2010). Nonverbal communication and detection of individual differences in social value orientation. In R. M. Kramer, A. E. Tenbrunsel & M. H. Bazerman (Eds.), *Social decision making: Social dilemmas, social values, and ethical judgments*. (pp. 147-170). New York, NY, US: Routledge/Taylor.

Smeesters, D., Wheeler, S. C., & Kay, A. C. (2009a). The role of interpersonal perceptions in the prime-to-behavior pathway. *Journal of Personality and Social Psychology, 96*(2), 395-414.

Smeesters, D., Yzerbyt, V. Y., Corneille, O., & Warlop, L. (2009b). When do primes prime? The moderating role of the self-concept in individuals' susceptibility to priming effects on social behavior. *Journal of Experimental Social Psychology, 45*(1), 211-216.

Sugden, R. (1984). Reciprocity: The supply of public goods through voluntary contributions. *Economics Journal, 94,* 772–787.

Suleiman, R., Budescu, D. V., Fischer, I., & Messick, D. M. (2004). Contemporary psychological research on social dilemmas. (pp. 315-331). New York, NY, US: Cambridge University Press.

Takahashi, C., Yamagishi, T., Tanida, S., Kiyonari, T., & Kanazawa, S. (2006). Attractiveness and cooperation in social exchange. *Evolutionary Psychology, 4*, 315-329.

Tazelaar, M. J. A., Van Lange, P. A. M., & Ouwerkerk, J. W. (2004). How to cope with "noise" in social dilemmas: The benefits of communication. *Journal of Personality and Social Psychology, 87(6)*, 845-859.

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly, 44*(4), 684-707.

Tenbrunsel, A. E., & Northcraft, G. (2010). In the eye of the beholder: Payoff structures and decision frames in social dilemmas. In R. M. Kramer, A. E. Tenbrunsel & M. H. Bazerman (Eds.), *Social decision making: Social dilemmas, social values, and ethical judgments*. (pp. 95-116). New York, NY, US: Routledge/Taylor & Francis Group.

Utz, S. (2004a). Self-activation is a two-edged sword: The effects of I primes on cooperation. *Journal of Experimental Social Psychology, 40*(6), 769-776.

Utz, S. (2004b). Self-construal and cooperation: Is the interdependent self more cooperative than the independent self? Self and Identity, 3(3), 177-190.

Utz, S., Ouwerkerk, J. W., & Van Lange, P. A. M. (2004). What is smart in a social dilemma? differential effects of priming competence on cooperation. *European Journal of Social Psychology, 34*(3), 317-332.

Utz, S., Van Lange, P. A. M., Green, E., Waldzus, S., & Bovina, I. (2005). Mary honest always friendly and is: Can scrambled sentences enhance prosocial motivation? *Representative Research in Social Psychology, 28*, 59-70.

van Dijk, E., de Kwaadsteniet, E. W., & De Cremer, D. (2009). Tacit coordination in social dilemmas: The importance of having a common understanding. *Journal of Personality and Social Psychology, 96*(3), 665-678.

van Dijk, E., Wilke, H., Wilke, M., & Metman, L. (1999). What information do we use in social dilemmas? Environmental uncertainty and the employment of coordination rules. *Journal of Experimental Social Psychology, 35*(2), 109-135.

van Dijk, E., Wit, A. P., Wilke, H. A. M., & de Kwaadsteniet, E. W. (2010). On the importance of equality in social dilemmas. In R. M. Kramer, A. E. Tenbrunsel & M. H. Bazerman (Eds.), *Social decision making: Social dilemmas, social values, and ethical judgments.* (pp. 47-69). New York, NY, US: Routledge/Taylor & Francis Group.

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., & Joireman, J. A. (2010). Social and temporal orientations in social dilemmas. In R. M. Kramer, A. E. Tenbrunsel & M. H. Bazerman (Eds.), *Social decision making: Social dilemmas, social values, and ethical judgments.* (pp. 71-94). New York, NY, US: Routledge/Taylor & Francis Group.

van Lange, P. A. M., & Kuhlman, D. M. (1994). Social value orientations and impressions of partner's honesty and intelligence: A test of the might versus morality effect. *Journal of Personality and Social Psychology, 67*(1), 126-141.

van Lange, P. A. M., Ouwerkerk, J. W., & Tazelaar, M. J. A. (2002). How to overcome the detrimental effects of noise in social interaction: The benefits of generosity. *Journal of Personality and Social Psychology, 82(5)*, 768-780.

van Lange, P. A. M., & Visser, K. (1999). Locomotion in social dilemmas: How people adapt to cooperative, tit-for-tat, and noncooperative partners. *Journal of Personality and Social Psychology, 77*(4), 762-773.

van Vugt, M., & De Cremer, D. (1999). Leadership in social dilemmas: The effects of group identification on collective actions to provide public goods. *Journal of Personality and Social Psychology, 76*(4), 587-599.

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: The sequel. A kernel of truth in predictive cheating detection. *Evolution and Human Behavior, 28*(4), 260-271.

Weber, J. M., Kopelman, S., & Messick, D. M. (2004). A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review, 8*(3), 281-307.

Wildschut, T., & Insko, C. A. (2007). Explanations of interindividual—intergroup discontinuity: A review of the evidence. *European Review of Social Psychology, 18*, 175-211.

Wit, A., & Wilke, H. A. (1990). The presentation of rewards and punishments in a simulated social dilemma. *Social Behaviour. Special Issue: Applied Perspectives on Social and Temporal Dilemmas, 5*(4), 231-245.

Wit, A., & Wilke, H. (1998). Public good provision under environmental social uncertainty. *European Journal of Social Psychology, 28*(2), 249-256.

Wubben, M. J. J., De Cremer, D., & van Dijk, E. (2009). How emotion communication guides reciprocity: Establishing cooperation through disappointment and anger. *Journal of Experimental Social Psychology, 45*(4), 987-990.

Wubben, M. J. J., De Cremer, D., & van Dijk, E. (2008). When emotions of others affect decisions in public good dilemmas: An instrumental view. *European Journal of Social Psychology, 38*(5), 823-835.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology, 51*(1), 110-116.

Yamagishi, T. (1992). Group size and the provision of a sanctioning system in a social dilemma. In W. B. G. Liebrand, D. M. Messick & H. A. M. Wilke (Eds.), *Social dilemmas: Theoretical issues and research findings*. (pp. 267-287). Elmsford, NY, US: Pergamon Press.

Yamagishi, T. (2001). Trust as a form of social intelligence. In K. S. Cook (Ed.), *Trust in society.* (pp. 121-147). New York, NY, US: Russell Sage Foundation.

Yamagishi, T., & Hayashi, N. (1996). Selective play: Social embededness of social dilemmas. In W. Liebrand & D. Messick (Eds.), *Frontiers of social dilemma research* (pp. 363-384). Berlin: Springer-Verlag.

Yamagishi, T., Kanazawa, S., Mashima, R., & Terai, S. (2005). Separating trust from cooperation in a dynamic relationship: Prisoner's dilemma with variable dependence. *Rationality and Society, 17*(3), 275-308.

Yamagishi, T., & Suzuki, N. (2010). An institutional approach to culture. In M. Schaller, A. Norenzayan, S. J. Heine, T. Yamagishi & T. Kameda (Eds.), *Evolution, culture, and the human mind.* (pp. 185-203). New York, NY, US: Psychology Press.

Yamagishi, T., Tanida, S., Mashima, R., Shimoma, E., & Kanazawa, S. (2003). You can judge a book by its cover: Evidence that cheaters may look different from cooperators. *Evolution and Human Behavior, 24*(4), 290-301.

Yu, C., Au, W., & Chan, K. K. (2009). Efficacy = endowment × efficiency: Revisiting efficacy and endowment effects in a public goods dilemma. *Journal of Personality and Social Psychology, 96*(1), 155-169.

Zaatari, D., & Trivers, R. (2007). Fluctuating asymmetry and behavior in the ultimatum game in Jamaica. *Evolution and Human Behavior, 28*(4), 223-227.