

Utilitarians aren't always fair & the fair aren't  
always utilitarian: Distinct motives for  
cooperation<sup>1</sup>

Lívia Markóczy  
Anderson Graduate School of Management  
University of California, Riverside  
Phone: +1 909 787-3908  
Fax: +1 909 787-3970  
Livia.Markoczy@ucr.edu  
<http://www.goldmark.org/livia/>

Version 1.78  
September 1, 2004

<sup>1</sup>The version you are reading is 1.78 last modified on September 1, 2004. Please quote that in any comments. This is a *draft* version. I am grateful for the support of Anderson Graduate School of Management, Southern California Edison, and Riverside Public Utility Company for financing and supporting the data collection and access. I would also like to thank Katherine Randazzo for her assistance in preparing the data for analysis and Martena Wilson for her essential help in bringing all the parties together and Jeffrey Goldberg for his suggestions and comments throughout this project. I would also like to thank Rachel Croson, Robyn Dawes, John Haleblian, David Mayers, Kathleen Montgomery and Dan Ozer for their helpful comments on earlier versions of this paper. Some draft of this paper and related material can be found at <http://www.goldmark.org/livia/papers/energy/>

## **Abstract**

I discuss fairness as a motive for cooperation in a social dilemma and investigate how various situational factors affect this motive in comparison to another cooperative motive: utilitarianism. I do this in the context of a large scale social dilemma: the California electricity crisis (2000–2001). Fairness motive refers to trying to reach a fair outcome by doing one's share but by not doing more than others do. Utilitarianism is a preference to maximize the collective outcome. I show that – although both of these motives can lead to cooperative behavior in a one-shot social dilemma – the behavioral consequences of these motives are affected differently by situational factors, such as perceived criticality of a situation, ownership of a problem, efficacy beliefs, and expectations of others' behavior. Utilitarianism and a fairness motive are captured by a survey of 700 consumers in California while cooperative behavior is calculated from actual electricity consumption data acquired from California utility companies.

Why do some people cooperate in situations in which it is against their narrow self-interest to do so, and which would lead to a collective disaster if nobody cooperates? This question has long interested social scientists because these types of situations are common within groups, organizations, and in societies; and the effective functioning and the survival of these collectives depends on the willingness of individuals to cooperate, even if it is not in their narrow self-interest to do so (Kramer, Brewer, and Hanna, 1996; Pruitt and Kimmel, 1977). For example, each individual is better off enjoying the convenience of driving their car to work as opposed to using public transportation, but, if everybody follows their narrow self-interest, this will lead to traffic jams and pollution on the collective level. Situations like these, in which a self-interest maximizing behavior of an individual is in conflict with maximizing collective welfare belong to class of problems that scholars call social dilemmas (Messick and Brewer, 1983).

Social dilemmas include two specific classes of dilemmas: “resource dilemmas” and “public good problems”. Resource dilemmas which Hardin (1968) calls the “tragedy of commons” occur when individuals decide how much to take from a shared resource (e.g., water). In a public good problem, there is some public good, say a lighthouse, which all will benefit from even if not everyone contributes to it, but the community is better off if everyone contributes. An individual is best off if everyone else contributes, but the individual doesn’t. An individual is worst off if that individual contributes and nobody else does.

I use the word “cooperation” to mean the behavior of contributing in a public good problem or exercising restraint in harvesting in a resource dilemma. I use the word “defection” to be the behavior of not cooperating.

Research on social dilemmas has investigated many explanations for why some individuals choose to go against their narrow self-interest by cooperating in such situations. One of these explanations centers around differences in individual motives that affect an individual’s cooperativeness. Motives are defined here as relatively stable individual tendencies to strive to approach a certain class of positive goals or to avoid some negative consequences or threats (Atkinson, 1982). This definition of a motive is broader than used in social dilemma research that traditionally limits attention to so-called social motives which are defined in terms of individual preferences for different distributions of outcomes to self and others (Messick and McClintock, 1968).<sup>1</sup>

---

<sup>1</sup>The difference between the two notions is that one is restricted only to considering the

Although a large number of social motives can be distinguished in principle, most empirical and theoretical attention has been given to three distribution preferences: (1) a preference to maximize the outcome of the collective as a goal (utilitarianism); (2) a preference to maximize the outcome for the self as a goal (greed or individualism); and (3) a preference to maximize the difference between the self and others as a goal (competitive motive or spite). Not surprisingly, a consistent finding across a large number of social dilemma studies is that utilitarians are most likely to cooperate across a wide variety of situations relative to those who are driven by greed or by competitiveness (e.g, Kuhlman, Brown, and Teta, 1992; Kuhlman, Camac, and Cunha, 1986b).

Although the role of an utilitarianism in leading to cooperation in a social dilemma has been widely explored, less attention has been given to another motive that may also lead to cooperation in the face of self-interest, the fairness motive.<sup>2</sup> Fairness as a motive is a desire to reach a fair outcome by simultaneously wanting to do one's fair share but also not wanting to contribute more than others do. (What is perceived as a fair outcome will be discussed below.) Although both utilitarianism and fairness may lead to cooperation, few studies have tried to elaborate the meaning of fairness as a motive or to look at the behavioral effects of these motives in the same social dilemma. Thus, little is known about whether these motives are affected similarly or differently by various aspects of a situation leading to different behavioral outcomes. In other words, we do not know under what conditions one of these social motives is more likely to lead to cooperation than another. The goal of this study is to contribute to our understanding of this issue by doing two things: (1) by elaborating on the concept of fairness as a motive; and (2) by hypothesizing what contextual factors are likely to lead to different behavioral outcomes between an utilitarianism and a fairness motive. I test these hypotheses in the context of a real life social dilemma, the California

---

aspects of the distribution of pay-offs. The other may, for example, include consideration of endowment and other facts about the situation that aren't represented directly in the pay-offs.

<sup>2</sup>Note that there is a substantial body of literature that is about how the perception of fair treatment affects people's cooperative behavior (see Colquitt, Conlon, Wesson, Porter, and Ng (2001) for an overview). But as I make clear later in the paper, a fairness motive is about a preference to behave in a particular way, which is distinct (although a related concept) from a preference for fair treatment. Fairness as a motive has been incorporated into a more general "prosocial" motive (e.g. Van Lange, 1999), but has rarely been addressed separately as a motive.

electricity crisis during 2000–2001, with the participation of 700 consumers.

## 1 Fairness

In this paper I take several atypical approaches to fairness. First, for this study it is important to cast fairness as a motive. Second, I do this while abstracting away from the distinctions among various fairness rules (such as equity and equality of outcomes, etc), while still maintaining these important distinctions. Third, I decompose fairness into two submotives: a desire to contribute no less than others and a desire to contribute no more than others.

### 1.1 Fairness as a motive

I treat fairness as motive in a narrow sense, meaning to have a desire to achieve particular goals or avoid particular outcomes. The details of the particular outcomes in the case of fairness will be described later. Treating fairness as a motive is distinct from how fairness is typically treated in the literature, which is to use it to understand how people respond to (un)fairness. For example, research has addressed the questions of why people care about whether they are treated fairly by others and of how people’s *judgment* of whether they are treated fairly in terms of procedures or in terms of the division of outcomes affects their behavior (e.g., Ambrose, 2002; Cohen-Charash and Spector, 2001; Lind and Tyler, 1988; Miles, Hatfield, and Huseman, 1989). But a strong fairness response and a strong fairness motive may not always co-exist in an individual. It is easy to imagine individuals who are highly sensitive to being treated unfairly, yet are not motivated to behave fairly themselves. Indeed Messick and Sentis (1979) have shown that there is a distinction between fairness judgments and fairness behavioral choice.

A motive to behave fairly may be rooted in a desire to behave morally: “to do the right thing” (Ambrose, 2002). For example, Lerner (1977) suggests that people are often motivated to be fair in their actions. Individuals may also prefer to behave fairly not because they see some intrinsic value to a fair outcome, but rather to avoid some negative social consequences that could be expected when someone behaves unfairly (Elster, 1989). They may also just use fairness as a simple heuristic such as: “contribute equally” or “harvest equally” thereby minimizing the mental computation of how much to contribute (Allison and Messick, 1990). For the purposes of this discussion,

a willingness to behave in certain ways to achieve a fair outcome is a fairness motive, irrespective of the various reasons why one might seek a fair outcome. Of course it must also be noted that what counts as a fair outcome is far from simple and will be discussed below.

Throughout the rest of this paper, my references to fairness will be about fairness as a motive, unless explicitly indicating otherwise. Now that we have distinguished, at least conceptually, fairness as a motive we need to elaborate on what sorts of outcomes are sought by a fairness motive.

## 1.2 Abstracting fairness

When scholars talk about various kinds of “fairness rules,” such as equality of inputs, equality of outcomes, equity, and need-based fairness, it is no coincidence that we call all of these “fairness.” The distinctions among these are important and they operate in systematically different ways (e.g., Adams, 1965; Deutsch, 1975, 1985; Greenberg, 1978; Hegtvedt and Cook, 2001; Messick and Cook, 1983), but there is also, at a more abstract level, something that unifies all of these under a common notion of fairness. The distinction that I wish to make in the following section should apply across the board to all of these fairness rules, so here I elaborate on a more abstract notion of fairness that covers all of the rules.

Messick (1980) has already done most of the work of the required abstractions. He argues that all of the fairness rules can be cast as seeking equality of something. The rules merely vary on what that something is. Below I provide examples for how various fairness rules can be recast this way.

Using  $I_s$  to represent one’s own input and  $I_o$  to represent the average of others’ inputs, input-based fairness can be represented by a desire to achieve the equality in equation 1.

$$I_s = I_o \tag{1}$$

Likewise, with  $O_s$  and  $O_o$  as outputs of self and others respectively, output-based fairness can be represented as a desire to achieve the equality in equation 2.

$$O_s = O_o \tag{2}$$

If we treat  $E$  as prior endowment, then need based fairness can be seen as trying to achieve the equality in equation 3.

$$O_s + E_s = O_o + E_o \tag{3}$$

Equity can be seen as a desire to achieve equal ratios of inputs and outputs as in equation 4

$$\frac{O_s - I_s}{I_s} = \frac{O_o - I_o}{I_o} \quad (4)$$

Likewise we can have need based equity, a desire to move toward the equality in equation 5

$$\frac{O_s + E_s}{I_s} = \frac{O_o + E_o}{I_o} \quad (5)$$

Other examples of fairness rules have been proposed, but for present purposes I only need to use a few such rules to illustrate the generalization I propose.

To abstract away from the specifics of each equality, I will present a meta variable  $\mathbf{R}$  for the results of some transaction or interaction.  $\mathbf{R}_s$  scopes over the left hand side of the equalities in equations 1–5, and  $\mathbf{R}_o$  scopes over the right hand side of those equalities. Thus, on this equality-based view of fairness, fairness is a desire to achieve the equality in equation 6

$${}_i\mathbf{R}_s = {}_i\mathbf{R}_o \quad (6)$$

where the  $i$  subscript indexes a fairness rule.

I will depart from the equality-based view of fairness in the following section, but I have used it here both to lay a foundation for what is to come and to illustrate that we can abstract away from particular fairness rules while remaining able to distinguish among them.

Following Van Lange (1999) one reason to depart from the equality-based view is that it fails to capture the intuition that people aren't so much seeking equality as seeking to minimize inequality. Mathematically, seeking equality and seeking to avoid inequality are the same, but psychologically they are distinct. We are already familiar with the fact that mathematically identical situations can be psychologically distinct if framed differently. Thus, it is useful to present the above equalities as instead an attempt to minimize inequality. In the minimize inequality view we represent the different fairness rules as trying to achieve:

$$0 = |I_s - I_o| \quad (7a)$$

$$0 = |O_s - O_o| \quad (7b)$$

$$0 = |(E_s + O_s) - (E_o + O_o)| \quad (7c)$$

$$0 = \left| \frac{O_s - I_s}{I_s} - \frac{O_s - I_o}{I_o} \right| \quad (7d)$$

$$0 = \left| \frac{E_s + O_s}{I_s} - \frac{E_o + O_o}{I_o} \right| \quad (7e)$$

And the generalized form of the minimize inequality view of fairness is as an attempt to achieve equation 8

$$0 = |{}_i\mathbf{R}_s - {}_i\mathbf{R}_o| \quad (8)$$

While I agree with Van Lange (1999) that minimizing inequality better represents the motives than seeking equality does, I do not feel that it goes far enough. In the next section I develop this further.

### 1.3 Two submotives of fairness

There are two sides to fairness. One side is a desire to not do less than (or receive no more than) others. That I call the contributory side of fairness (c-fair). The detracting side of fairness (d-fair) is the desire to not contribute more than (or receive less than) others. D-fair can be seen as a desire to avoid the indignance of others free-riding while you contribute, and c-fair can be seen as a desire to avoid the shame of free-riding while others contribute. Using the notation of the previous section, the generalized view of fairness is to achieve

$${}_i\mathbf{R}_s \not\prec {}_i\mathbf{R}_o \quad (\text{c-fair})$$

$${}_j\mathbf{R}_s \not\prec {}_j\mathbf{R}_o \quad (\text{d-fair})$$

This provides us with two distinct submotives of fairness. C-fair is to not end up better than others (where “better” depends on the particular fairness rules being attended to in a given situation), and d-fair is a motive to not end up worse than others (where “worse” depends on the particular fairness rules in question). Note that c-fair is expressed as above, with the relation

$\not\leq$ , and not as the mathematically equivalent  ${}_i\mathbf{R}_s \leq {}_i\mathbf{R}_o$ . This is to represent that people aren't so much trying to achieve a particular state, but trying to avoid its opposite.

This formulation is mathematically equivalent to the “minimize inequality” view, which in turn is mathematically equivalent to the “seek equality” view. That is, the solution to these inequalities is to minimize the absolute difference between  $\mathbf{R}_s$  and  $\mathbf{R}_o$  and the solution to minimizing that difference is to seek equality. However, the c-fair and d-fair conception may better capture what is actually motivating people. Just as indignance and shame are distinct emotions, the attempts to avoid them should also be distinguished.

What is striking about the view I put forward here is that it is widespread in the informal discussions of fairness in most of the literature. A scholar discussing the issue will provide examples to illustrate both d-fair and c-fair and describe them in different terms, but when it comes time to actually define fairness, it will be defined in terms of seeking equality or minimizing inequality. Examples of this abound. Messick and Sentis (1979, p. 419) provide an example:

Adams [1965] asserts that inequity is aversive, creating dissatisfaction if one's outcome to input ratio is smaller than the other's, and creating feelings of guilt if one's ratio is larger than the other's. It is further proposed that less inequity is preferred to more, i.e., that smaller differences between the ratios is preferred to larger differences.

Here we see that Messick and Sentis (paraphrasing Adams) describe two separate reactions, anger and guilt, to two distinct situations. That is, they initially make the distinction between d-fair and c-fair. But also note that they further describe this as a single thing, a dispreference for large differences in the ratios.

Other considerations are possible in making the c-fair/d-fair distinction that cannot be expressed with the other views. The first is the possibility that in a specific situation an individual might use one fairness rule for c-fair and another fairness rule for d-fair. While I have no specific reason to believe that that happens, it would not be a good idea to rule out that possibility before even considering it. More importantly, by breaking up fairness into these two motives we can acknowledge that people may give more weight to c-fair than to d-fair or may give more weight to d-fair than to c-fair. The minimize inequality view – based as it is at looking at an absolute difference

– is unable to express that possibility. The results of the study I report in this article strongly suggest that, while there is a correlation between the strength of the c-fair and d-fair motives in an individual, they can and do go their separate ways.

C-fair and d-fair require attending to information about what others do or receive. And as with all such attention, it may be subject to the confirmatory bias. If someone suspects that they are failing to achieve c-fairness, they will seek out evidence which will support that view and may not notice evidence to the contrary. If someone suspects that are not achieving d-fairness, they may seek out evidence which confirms that suspicion. While at this point I can provide no evidence that the confirmatory bias plays a role in fairness this way, this plausible hypothesis cannot even be considered without separating c-fair from d-fair.

## 1.4 The fair duality

I have argued that fairness ought to be considered as composed of two sub-motives, c-fair and d-fair. But the question of what “submotive” means remains. Should we abandon a unified notion of fairness, or does it remain meaningful to talk about fairness motivated individuals? I argue that c-fair and d-fair, which are very important to distinguish, comprise a more general fairness motive. My argument is based on the fact that in both the present study and in a pilot study (involving an artificial social dilemma and 200 MBA students at a British university) c-fair and d-fair were the only pair of cooperative and non-cooperative motives to positively correlate with each other. Other pairs of motives did positively correlate, but this was the only pair of such motives where one was a cooperative motive and the other a motive for defection. Typically, cooperative motives correlated *negatively* with non-cooperative motives.<sup>3</sup> That c-fair and d-fair positively correlate under these circumstances strongly indicates that there is a single and coherent notion of fairness that contains these two submotives.

Broadly speaking we can talk about a fairness motive, and we can assert that people can vary in the degree to which they are motivated by fairness.

---

<sup>3</sup>In (self-citation) I describe why my data may turn this up even if people are disposed to have a mixture of both cooperate and non-cooperative motives. In short the answer is that despite attempts to compensate in the wording of the questionnaire, there will remain a tendency of subjects to mostly highlight those motives that ended up supporting their actual decision in the social-dilemma.

At the same time we can maintain that when measuring fairness or examining its role we must distinguish between c-fair and d-fair.

If fairness has these two sides then it can lead both to cooperative behavior through c-fair, and it can also lead to defection through d-fair. All of this is sensitive to the perceived or expected behavior of others. Because fairness leads to either cooperation or defection, the theory here can make no (simple) prediction about whether fairness-motivated people are more cooperative or less cooperative than others.

## **2 Utilitarianism as a motive**

Utilitarianism, just like fairness, is another motive that may lead to cooperation in social dilemmas in the face of short-term self-interest. Since utilitarianism has received much attention in the literature (e.g, Kuhlman et al., 1992, 1986b; Liebrand, 1984) I only provide here a short overview of this concept. Utilitarianism (also labeled “cooperative motive” (e.g. Kuhlman, Camac, and Cunha, 1986a; Kuhlman and Marshello, 1975; MacCrimmon and Messick, 1976) or “pro-social motive”<sup>4</sup>) is a relatively stable individual tendency to try to maximize the benefit for the entire collective (including oneself).

There are many reasons why people may prefer to place the interest of the collective above self-interest. People may feel empathy towards others (Batson and Shaw, 1991), or utilitarian preferences may have an origin in human evolution (Hoffman, 1981; Reeve, 1998; Simon, 1990, 1993), or simply because they feel that by behaving as a utilitarian they may be able to induce others to behave the same way towards them (e.g., self-citation).

### **2.1 Utilitarianism isn’t altruism, except for when it is**

There is an important distinction between utilitarianism and altruism, but as will be seen later, I also blur that distinction for much of the current study. As described in this section, there are important senses and circumstances under which utilitarianism and altruism can be treated as part of the same concept.

---

<sup>4</sup>I chose not to use these labels to avoid confusion between cooperative motive and the act of cooperation and because pro-social motive is often defined in a way that it includes fairness beside utilitarianism (e.g., Van Lange, 1999).

Motive	Weight of $u_s$	Description
Greed	$w_s = 1$	Maximize own utility
Utilitarianism	$w_s = \sqrt{1/(n+1)}$	Maximize total utility
Altruism	$w_s = 0$	Maximize others' utility

Table 1: Weights and Motives: Several motives expressed as maximizing equation 9 with various values for  $w_s$

Utilitarianism tends to be distinguished from altruism in the social dilemma literature (e.g. Kuhlman et al., 1986b; MacCrimmon and Messick, 1976). Altruism is defined as a desire to maximize the outcome of others' while being indifferent to the outcome to the self; utilitarianism is defined as trying to maximize a joint outcome that includes the self. Here I argue that distinguishing utilitarianism from altruism is not always meaningful. To show this, I will recast the representations of social motives in a way that will help make it clear that, although there is an important difference between altruism and utilitarianism, there are also many situations in which an orientation toward one will be an orientation toward the other.

Altruism and several other motives can be expressed in terms of the weights we assign to our own utility,  $w_s$ , and to other individuals,  $w_i$  where  $i$  indexes the  $i$ -th non-self individual. In these formulations, people attempt to maximize the expression in 9.

$$w_s u_s + \sum_{i=1}^n w_i u_i \tag{9}$$

where  $u_i$  is the utility of the  $i$ th non-self individual and  $n$  is the number of non-self individuals. The sum of all of the squares of the weights must equal 1 (Liebrand, 1984). That is,  $w_s^2 + \sum_{i=1}^n w_i^2 = 1$ .

Pure altruism would be when  $w_s = 0$  and  $\sum w_i^2 = 1$ . Pure greed would be when  $w_s = 1$  and  $\sum w_i = 0$ . Utilitarianism can be represented in this scheme as well; it is when  $w_s = \sqrt{1/(n+1)}$ . This is summarized in table 1.

In this view, altruism, utilitarianism and greed aren't so much distinct motives as merely different points on the same continuum. Also note that where  $n$  is very large, say the number of electricity consumers in California, utilitarianism becomes practically identical to altruism. It is very reasonable to expect that where  $n$  is large the vast majority of individuals will

fall between utilitarianism and greed. In such a situation, a measure of a tendency toward utilitarianism and a tendency toward altruism will be the same. Because this is a study of a very large scale social dilemma, I will use the terms utilitarianism and altruism interchangeably except when I need to distinguish between them. More importantly, there is no need to separate measures of altruism and utilitarianism in this context.

## 2.2 Behavioral consequences of utilitarianism

It has been consistently demonstrated that utilitarianism increases the likelihood of cooperation in a wide variety of contexts (e.g., Kuhlman et al., 1986b). Individuals with utilitarianism have been shown, for example, to be more likely to volunteer or to make a charitable contribution than individuals with other social motives do (e.g., Smith, Kehoe, and Cremer, 1995; Unger, 1991), or behave like good citizens in the workplace by helping others or by doing more for their organizations than is recognized in the formal reward system (e.g., Organ, 1988). That utilitarianism increases the likelihood of cooperation has been also widely shown in various social dilemmas (e.g., Kuhlman et al., 1992, 1986b; Liebrand, 1984). Thus I hypothesize that

**Hypothesis 1** *The more individuals are motivated by utilitarianism, the more likely they are to behave cooperatively.*

## 3 Interaction between motives and situation

Although both fairness and utilitarianism can lead to cooperation in a social dilemma, we know little about when these two motives lead to different behavioral outcomes because a fairness motive and utilitarianism have been rarely investigated in the same social dilemma. Given that these motives relate to the pursuit of different preferences, it is possible to imagine conditions that are helpful in realizing some preferences but not of others. Such conditions are likely to affect the behavioral outcomes of fairness differently than the behavioral outcomes of utilitarianism. In this paper I consider four aspects of situations in which I expect that utilitarianism would lead to a different behavioral outcome than a fairness motive would and will investigate the interaction of these motives with aspects of the situation. The situational aspects include: (1) the criticality of the public good problem,

meaning how substantial are the consequences of failing to achieve the public good; (2) the sense of ownership, meaning whether an individual perceives the problem as his or her problem; (3) the expected contribution of others in the situation; and (4) the expected effectiveness of one's own contribution in the given situation.

Individuals do not respond so much to a situation but to their perception of the situation. This distinction is rarely made when scholars look at the interaction between a person and a situation (e.g., Chatman and Barsade, 1995). Perceptions of the same situation may vary substantially among individuals, which points to the importance of distinguishing a situation from how it is perceived (Ross and Nisbett, 1991). Perception here means both the aspects of a situation that an individual notices, as well as the interpretation an individual assigns to these. In (self-citation) I elaborate my view about the relationship between situation, perception, multiple motives and behavior.

### 3.1 Motives and criticality of the situation

In the context of a public good problem, a situation is considered critical if the consequences of failing to achieve the public good are perceived to be substantial. In the context of resource dilemmas, a situation is considered critical if the consequences of resource depletion are perceived to be substantial.

I expect that perceived criticality of the situation will more strongly effect the utilitarian motive than the fairness motive. Recall that utilitarians are concerned with the collective benefit and so if the perceived benefit of a public good is very very good, then utilitarians will be more motivated to achieve that good. Fairness-motivated people, who are not concerned about actual benefit gained by the collective action, will be less sensitive to the actual value of the public good.

**Hypothesis 2** *Perceived criticality of the problem will more strongly (and positively) moderate the effect on cooperation of the utilitarian motive than of the fairness motive.*

### 3.2 Motives and ownership of a problem

In addition to perceiving the situation as critical, a person's willingness to assume responsibility has also been found to play a role in an individual's willingness to contribute to the common good (Schroeder, Sibicky, and Iris, 1995). The more an individual is willing to assume responsibility for the problem, the more likely it is that the person will cooperate (Schwartz and Gottlieb, 1980).

In the words of the fictional character Maxwell Smart, "it's not whether you win or lose, but how you lay the blame." That does not describe the utilitarian, where blame is assigned has little effect on collective outcome. The fairness-motivated person, on the other hand, may be highly sensitive to where the blame can be assigned, since they may feel that it is "only fair" for people to take care of the problems that they helped create (c-fair) and unfair to suffer the consequences of a problem that others created (d-fair). For a utilitarian, it is not relevant whether they are faced with "somebody else's problem," since the collective gain from cooperation is the same irrespective of ownership of the problem. Thus, I predict that

**Hypothesis 3** *Perceived ownership of a problem will more strongly (and positively) moderate the effect on cooperation of the fairness motive than of the utilitarian motive.*

### 3.3 Motives and expectations of others' cooperation

A cooperative solution can only emerge in most social dilemmas if a sufficient number of individuals are willing to cooperate. Thus, previous research predicted that individuals' choice to cooperate is affected by their expectations of whether others will also cooperate (e.g., Pruitt and Kimmel, 1977). Yet the expectations about the contribution of others are likely to affect the behaviors of utilitarians differently than the behavior of those with a fairness motive.

There are various explanations for why individuals make their cooperation dependent on what they expect others to do. Individuals may infer social norms from the expected behavior of others and thus follow this norm. The inclination of individuals to conform with others may lead them to follow what others are doing (Van Vugt, Meertens, and Van Lange, 1994). This is a form of fairness, because those who are fairness motivated are interested in matching their contributions with others. If others are not cooperating, they

are not obliged to cooperate either. Thus, higher expectations of others' cooperation are more to enhance the cooperation coming from a fairness motivated.

The moderating effect of expectations on utilitarianism is less clear, however. Recall that utilitarianism leads to cooperation in order to achieve a common good. Achievement of this goal is not problematic if enough others are expected to be cooperating; thus, a utilitarian does not need to put much effort into achieving this goal. On the other hand, if not enough others are expected to be cooperating, achievement of this goal becomes problematic and thus more cooperation is required from a utilitarian. This argument is consistent with Croson (1998), who argues that utilitarians increase their effort when others are not cooperating to compensate for the missing cooperation from others. Contrary to this view, however, Dawes, McTavish, and Shaklee (1977) and others argue that higher expectations of others' cooperation are more likely to correspond to cooperation from a utilitarian because utilitarians tend to base their expectations of others on their own intentions to cooperate (Messé and Sivacek, 1979). I expect these two contradictory mechanisms to weaken each other, leading to a weak or no effect. Thus, I predict that

**Hypothesis 4** *The perception that others are cooperating will more strongly (and positively) moderate the effect on cooperation of the fairness motive than of the utilitarian motive.*

### 3.4 Motives and efficacy beliefs

It is not just social motives that affect the cooperativeness of individuals but, also whether they expect that their contributions will make a difference in resolving a given social dilemma. This is a form of efficacy belief as defined by Kerr (1992, p. 102): “a belief that one’s cooperative acts will have some impact on the collective welfare.” The word “efficacy” is frequently used with this meaning in social dilemma research, but this meaning is less well known to other organizational theorists who tend to use “efficacy” in the meaning provided by Bandura (1986, p. 391): “a judgment of one’s capacity to accomplish a certain level of performance.” But, as Kerr (1992) stated, in social dilemmas, it is usually less of an issue whether one is able to cooperate and more so whether one’s contribution makes any difference at all in resolving a social dilemma. I will therefore use efficacy as defined

by Kerr (1992).

Individuals differ in their general efficacy beliefs (trait differences), as well as in beliefs of how important their contributions are in a specific social dilemma (task-specific differences). These latter beliefs are likely to be affected by individual expectations of the contributions of others; thus, efficacy beliefs and expectations about the behavior of others are not entirely independent.

That individuals care about whether their contributions make a difference has been well demonstrated in social dilemma research (e.g., Kerr and Harris, 1996; Olson, 1965). For example, it has been shown that individuals are more willing to cooperate in a smaller group than in a larger one due to the declining efficacy of individual contributions in a larger groups (Olson, 1965). Also, other researchers have shown that when people are told that their contributions are critical for reaching some particular threshold, they tend to increase their cooperation (Dawes, Orbell, Simmons, and van de Kragt, 1986).

Again, I expect that the impact of efficacy beliefs will have different moderating effects on fairness based cooperation versus utilitarian based cooperation.

Cooperation from utilitarianism will be reinforced by a belief that cooperating will make a real difference in achieving a collective benefit. The fairness motive, on the other hand, is less likely to be affected by efficacy beliefs, because fairness-motivated people care less about achieving collective benefit and more about matching their contributions with the contributions of others. Efficacy, like utilitarianism, is about outcome. Thus, I hypothesize that

**Hypothesis 5** *A sense of self-efficacy will more strongly (and positively) moderate the effect on cooperation of the utilitarian motive than of the fairness motive.*

An overview of these hypotheses is shown in Figure 1.

## 4 Methods

To test the above hypotheses, I conducted a study in a large-scale social dilemma: in the context of the California Electricity Crisis during 2000–2001.

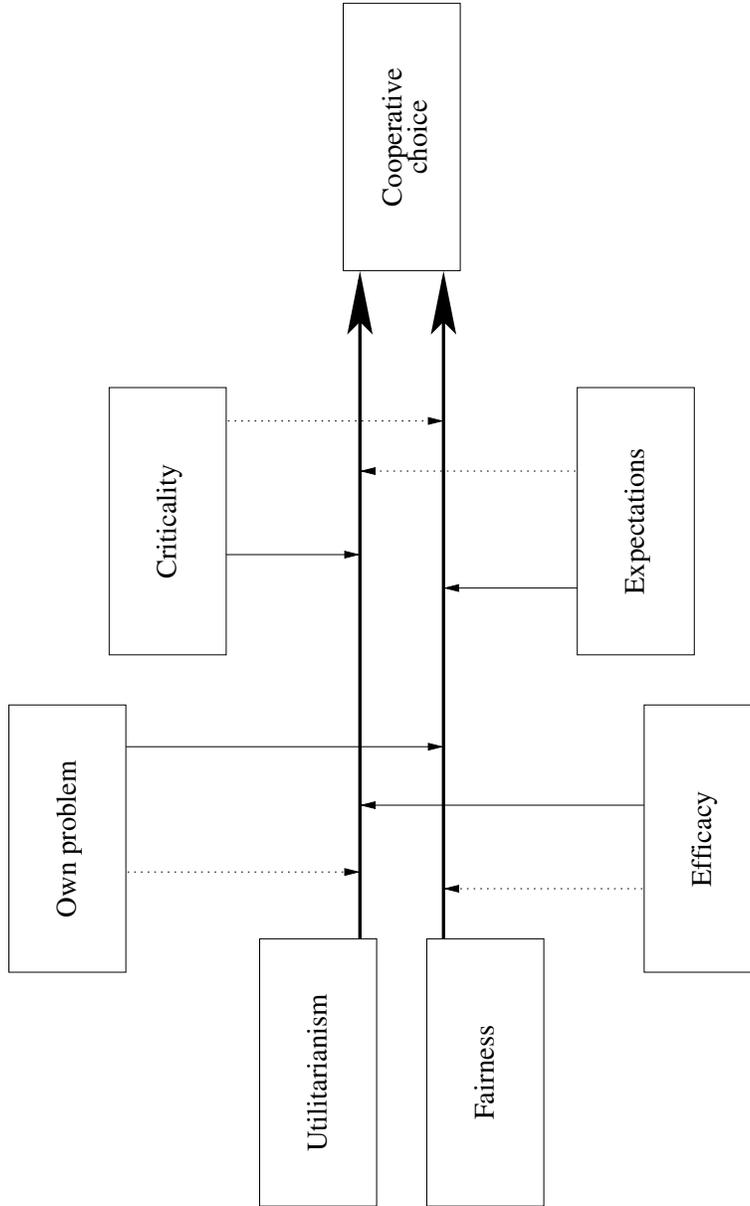


Figure 1: Overview of hypothesis: Dashed lines show weaker moderators while solid lines show strong ones.

During this time, California experienced electricity shortages for reasons that are beyond the scope of this paper (Borenstein, 2001).

What made the electricity crisis a social dilemma was that utility companies were not allowed by state regulation to raise the prices to consumers beyond a certain level. As a consequence of this, prices alone were not sufficient incentives for the demand to fall to the level of supply. Excess demand of electricity thus created a threat of black-outs, which occasionally led to a disruption in electricity supply to certain regions of California.

Because California was not in the position to increase the supply of electricity in the short term, the only way it could avoid the threat of black-outs (collective benefit) was to ask for a voluntary reduction of electricity usage by Californians beyond the cost-saving benefits that they gained from such behavior. This made the electricity crisis situation a large-scale, high-profile social dilemma. Although there was certainly individual conservation based on the price of electricity paid by consumers, the fact that people were being asked to conserve for the benefit of the whole state illustrates that this was, indeed, a social dilemma. It is not as clean and neat a social dilemma as a laboratory one, but we always have to sacrifice some tidiness of control when conducting a field study.

## 4.1 Sample

A representative sample of 11400 households was drawn from around 11 million households that are served by a municipal utility, Riverside Public Utility Company (RPUC) and by Southern California Edison (SCE). I selected these two utility companies because (1) while SCE raised the consumer prices as a response to the electricity crisis, RPUC did not raise its prices, which allows us to isolate the effect of price increase on electricity-saving behavior; and (2) RPUC and SCE serve households that fall into comparable climate conditions.

The representative sample was drawn based on five climate zones and on two housing types. These are the same climate zones used by the utility companies in their own modeling used for planning and operations. Housing types fall into two categories: multi-family residences that include apartments, condominiums, and town homes and single-family residences that include mobile homes. I excluded multi-family residences that do not use metering for individual households but use shared-metered accounts instead.

I sampled across climate zones to reach comparable samples from SCE

and RPUC in terms of weather conditions. I used housing type as a proxy measure for income, which tends to affect electricity consumption. Data were collected through telephone interviews by a major survey company and electricity consumption data for the sample were obtained from the utility companies. The company that conducted the telephone survey had been instructed to produce 700 completed questionnaires total<sup>5</sup> and they were also told how many completed questionnaires they needed to produce in each weather cell and in each housing type within these cells, so that the completed interviews would be a representative sub-sample of the total sample. The surveyors were also instructed to make sure that they interviewed individuals in the households who were the primary decision makers concerning the electricity saving behavior in the household. I used ANOVA to test for potential bias between the sample of 700 completed questionnaires and the rest of the full sample by comparing the electricity consumption in kilowatt-hours (kWh) of the respondents with the electricity consumption of the sample pool between June 1999 and September 2001. The result of the analysis revealed no significant difference in the electricity consumption between the respondents and the sample pool ( $F = .284, p = .594$ ).

## 5 Variables

The survey used to measure motives and the various situational factors was developed through the following steps: First, I used the survey in an exploratory study in the context of a social dilemma<sup>6</sup> with the participation of 191 MBA and Executive MBA students at a major UK university. In the exploratory study the following constructs were measured: (1) utilitarianism, (2) c-fair, (3) d-fair, (4) greed, (5) expectations of others' behavior; (6) and efficacy beliefs. I conducted a principal-component analysis with an oblique rotation (oblimin  $\delta = 0$ ) to test whether items loaded to factors as expected. Items loaded into the predicted six factors in the way that I expected, except that some items loaded into more than one factor with a factor loading .35 or higher. These items were eliminated from the final questionnaire. In

---

<sup>5</sup>The sample of 700 was decided based on two considerations: (1) to have a large enough sample to reach the statistical power that is needed to test the hypotheses; and (2) to stay within a given budgetary constraint. Also very substantial steps were taken to ensure the privacy of the individuals surveyed; I will provide details of that upon request.

<sup>6</sup>The text of this social dilemma is available from the author

the electricity study some items were slightly reworded to make them more applicable to the electricity crisis situation. For example, the item in the first questionnaire “I do not want to be a sucker” was changed to “I don’t want to be a sucker by conserving while others don’t.” In addition to minor rewording, I also added items to the questionnaires to measure criticality of the situation and ownership of the problem. The items used to measure utilitarianism, c-fair, d-fair and greed are listed in Table 2.<sup>7</sup> The items used to measure criticality of the situation, ownership of the problem, expectations of others’ behavior and efficacy beliefs are listed in Table 3. Prior to listing the items individuals were presented with the following question: *I’m going to read some statements that describe how you might think about whether or not to conserve. Please, tell me, on a scale of 1 to 5 how much each statement is like your own thinking about conservation. 1 means “not at all like my thinking” and 5 means “very much like my thinking” about electricity conservation.*

I conducted a confirmatory factor analysis for the electricity sample using EQS to test whether a correlated eight-factor structure (4 motives and 4 situational factors) fits the data. The overall fit of the eight factor structure was quite good. (The LISREL CFI=0.959, the LISREL AGFI=0.947, the RMSA=0.038 and the 90% confidence interval of RMSA is between 0.030 – 0.046). The correlations among the factors are listed in Table 4. The reliability of the measures is reported at the bottom of Table 2 and Table 3. Note that the reliability of the two items measuring perceived criticality is well below the accepted level of reliability, and thus findings that involve this measure need to be taken with a grain of salt.

I have argued throughout the paper that fairness has both a contributory (c-fair) and a detracting side (d-fair). Thus, I considered only those individuals motivated by fairness who scored high on *both* c-fair (4 or higher) and d-fair (4 or higher). These individuals are coded with the dummy variable 1. Those individuals who were high on neither c-fair (less than 4) nor d-fair (less than 4) were coded with a dummy variable 0. The remaining individuals (who were high on c-fair or d-fair but not both) were excluded from the analysis. This resulted in an effective sample for all fairness-related calculations of 575.

---

<sup>7</sup>Note, that I also measured greed with the survey, and thus I included this measure for the confirmatory analysis. This measure, however is not used in the study.

<b>Items</b>	<b>F 1</b>	<b>F 2</b>	<b>F 3</b>	<b>F 4</b>
	<b>Utilitarianism</b>	<b>Greed</b>	<b>D-fair</b>	<b>C-fair</b>
It's worth the effort to conserve if others benefit	.80			
It's worth a lot for me to know that I am making everyone in CA better off	.73			
Helping others is important	.71			
I like to help others out	.66			
By not conserving, I won't end up worse off than anyone else		.88		
If everyone else conserves a lot then I get the best of both worlds if I don't		.86		
If others put a large effort into conservation and I don't, that is my gain and their loss		.81		
Let the suckers put a lot of effort into electricity conservation		.63		
By not conserving I'll come out better off than those who conserve a lot		.59		
It is unfair to expect me to contribute more than others			.86	
I don't want to be a sucker by conserving while others don't			.85	
I don't want to do more than others			.78	
I don't want to risk putting in a large effort into electricity conservation when others may just make a small effort			.75	
If others conserve only a little why should I do more?			.73	
Everybody needs to share the burden if everybody wants to enjoy the benefit				.81
I should do my fair share				.79
If others put a lot of effort into conservation I should do the same				.56
Cronbach $\alpha$	.77	.76	.75	.70

Table 2: Results of confirmatory factor analysis

Items	F 1 Ownership	F 2 Efficacy	F 3 Expectations	F 4 Criticality
This is my problem as well as others'	.86			
The problem affects me, and I should put in some effort to fix it	.84			
Past practices of people like me created the problem, so I should help fix it	.76			
A small contribution goes a long way		.85		
My electricity conservation would be a big contribution to California		.83		
Lots of individual contributions add up to a big difference.		.64		
My fellow Californians will not put much effort into conserving*			.77	
I do not expect others to save much electricity*			.75	
Californians tend only think about themselves*			.72	
I believe that there is a real need for Californians to reduce electricity use				.69
I believe that there is a real shortage of electricity in California				.69
Cronbach $\alpha$	.77	.75	.70	.47

\*These items have been reverse coded

Table 3: Results of confirmatory factor analysis

	1	2	3	4	5	6	7	8
<b>1util</b>	1.00							
<b>2Greed</b>	-.24**	1.00						
<b>3d-fair</b>	-.19**	.36**	1.00					
<b>4c-fair</b>	.34**	-.19**	.17**	1.00				
<b>5Crisis</b>	.18**	-.42**	-.49**	.14**	1.00			
<b>6Own</b>	.45**	-.11**	-.14**	.32**	.28**	1.00		
<b>7Expect</b>	.22**	.02	-.22	-.06	.06	.09*	1.00	
<b>8Efficacy</b>	.33**	-.14**	.14**	.39**	.33**	.45**	.25**	1.00

† $p < .1$ ; \* $p < .05$ ; \*\* $p < .001$

Table 4: Correlations among the factors

## 5.1 Dependent variable

Electricity consumption of each household measured in June–August 2001 (the summer when Californians faced black-outs) was used as a dependent variable while electricity consumption in June–August 2000 (the summer when people were not yet broadly aware of the electricity crises) was controlled for. I reversed the sign of the dependent variable to reflect electricity saving behavior (cooperation) as opposed to electricity consumption.

## 5.2 Other Control variables

In addition to controlling for electricity consumption in the summer 2000, I also needed to control for situational and demographic differences among consumers that may affect electricity saving behavior.

SCE customers started out paying less for electricity than RPUC customers, but the SCE customers had a sharp rate increase in June 2001. To control for the effect of price increase for SCE consumers, I included a dummy variable marking whether a consumer is served by RPUC (1) or SCE (0).

Beyond their electricity providers, households differ in terms of how much heating and cooling is required. This was calculated by using the concept of “heating degree days” and “cooling degree days.” These are calculated by taking the average,  $\bar{T}$ , of the high and low temperatures (Fahrenheit) for a day in each climate zone. If  $\bar{T}$  is lower than 65°F then the number of heating degrees, HD, is taken as  $HD = 65 - \bar{T}$ , otherwise the heating degrees for the day is zero. When  $\bar{T}$  is greater than 74°F, for a particular day, the cooling degrees, CD, of that day is calculated as  $CD = \bar{T} - 74$ , and is zero otherwise.

The number of cooling degree days for a period (say a month) is the sum of the cooling degrees for each day in the period. Likewise the number of heating degree days in a period is the sum of the heating degrees for each day in that period. These measures are used by the utility companies and the California Energy Commission to forecast demand. To control for the weather, both in the summer 2000 and 2001, I used as a control variable the cooling degree days during the summer of 2000 and of 2001. Similarly, I used the heating degree days for the summer of 2000 and of 2001. These data for each climate zone were provided by the utility companies.

I included four additional demographic variables as controls, as these are likely to affect electricity consumption. These variables are all self-stated measures, including household income (1 for less than \$20,000; 2 for between \$20,000–\$40,000; 3 for between \$40,000–\$70,000; 4 for between \$70,000–\$100,000 5 for more than \$100,000), building type (coded as 1 for multiple family homes and 0 for single family homes), whether the property is owned or rented (coded as 1 for owned and 0 for rented) and the number of people living in the household.

## 6 Analysis

Hierarchical regression analysis was used to control for the effect of demographic and situational differences while assessing the effects of the main variables (utilitarianism and a fairness motive) and the moderator variables (criticality of the situation, ownership for the problem, expectations of others, and self-efficacy beliefs) on cooperation as well as the interaction effects of the motives and the moderator variables. In the first step, I entered the mean electricity consumption in June–August 2000 as the control variable. In the next step I entered the rest of the control variables in one block (mean cooling degree days in the summer 2001, mean heating degree days in the summer 2001, being a consumer that is served by SCE versus RPUC, number of people living in the household, income, own or rent, and building type). Next, I entered measures of criticality of the situation, ownership for the problem, expectations of others, self-efficacy beliefs, utilitarianism and fairness to assess the effects of the moderator variables and motives on cooperative behavior. Finally, I entered the interactions between motives and the moderator variables. I assessed the statistical significance of the block of variables entered at each step with an  $F$ -test of the change in  $R^2$ . Within

significant blocks of variables, I appraised the statistical significance of the effect of each variable with a  $t$ -test.

The regression provides only a guide to the overall effect of variables on the model. Hypotheses 2 through 5 make more specific predictions, so a more specific and powerful test is needed. We need to know if some moderator has a stronger interaction on one of the motives than on the other. Following Johnson, Johnson, and Buse (1987, p. 127), I tested this with the following statistic calculated in equation 10:

$$t_{\text{diff}} = \frac{\beta_1 - \beta_2}{\sqrt{\text{var } \beta_1 + \text{var } \beta_2 - 2 \text{cov}(\beta_1, \beta_2)}} \quad (10)$$

where  $\beta_1$  is the regression coefficient for the interaction between the moderator and the independent variable for which I expect the stronger interaction;  $\beta_2$  is the regression coefficient for the interaction between the moderator and the independent variable for which I expect the weaker interaction. Note that I am not taking the variance and covariance of the variables directly, but of the coefficients from the regression analysis. From the  $t_{\text{diff}}$ , probabilities are calculated using a one-tailed cumulative Student- $t$  distribution. The  $t$ -test is one tailed when we have a specific hypothesis (as we do in these cases) that one particular regression effect,  $\beta_1$ , is stronger than another particular effect,  $\beta_2$ . The  $t$ -test is calculated with the degrees of freedom,  $\nu$ , as  $\nu = n - (K + 1)$  where  $n$  is the sample size (575 in our case) and  $K$  is the number of independent and control variables in the regression (14), so  $\nu = 561$ .

## 7 Results

Cross-correlations among the control variables, the main variables and the moderators are shown in Table 5. The result of the hierarchical regression analysis is presented in Table 6.

In step 1 I entered the mean electricity consumption in June–August 2000. Not surprisingly, the effect of this variable on the electricity saving in June–August 2001 was highly, and negatively significant ( $t = -52.68$ ,  $p < .001$ ) explaining 85.4% of the variation. In step 2 I entered the rest of the control variables, including cooling degree days for the summer 2001, heating degree days for the summer 2001, whether households are served by RPUC or SCE, income, home ownership, building type, and number of people

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 <b>KWh01</b>	1.00													
2 <b>KWh00</b>	.92**	1.00												
3 <b>Cool01</b>	.32**	.28**	1.00											
4 <b>Heat01</b>	-.19**	-.18**	-.49**	1.00										
5 <b>RPUC</b>	-.03	.00	.15**	-.05	1.00									
6 <b>Income</b>	.02	.03	-.02	.06	-.05	1.00								
7 <b>Own</b>	.02	.02	.06	-.06	.04	.43**	1.00							
8 <b>Bldgtype</b>	-.03	-.01	-.01	.02	-.04	-.29**	-.36**	1.00						
9 <b>#inhouse</b>	.04	.00	-.00	.02	.01	-.26**	-.10*	.30**	1.00					
10 <b>Crisis</b>	-.02	-.01	.05	-.03	.08*	.13**	.06	-.00	-.07*	1.00				
11 <b>My-prob</b>	-.01	.00	.02	-.03	.02	.04	.05	-.04	-.05	.28**	1.00			
12 <b>Expect</b>	-.02	-.00	.01	-.07	.04	.05	.05	-.00	-.02	.06	.09*	1.00		
13 <b>Efficacy</b>	-.05	-.02	.02	.02	-.05	.18**	.03	-.05	-.09	.33**	.45**	.25**	1.00	
14 <b>Util</b>	-.04	-.01	.01	.05	.07	.00	.06	-.04	-.08	.17**	.43**	.21**	.43**	1.00
15 <b>Fair</b>	-.00	.00	.03	-.04	-.02	-.14**	-.01	.01	.02	-.31**	.18**	-.24**	.20**	.15**

†  $p < .1$ ; \*  $p < .05$ ; \*\*  $p < .001$

Table 5: Correlations

living in the house. The overall effect of these control variables on electricity saving behavior was significant ( $(\Delta F = 2.25, p < .05)$ ). Among the variables, the mean cooling degree days ( $t = -2.21, p < .01$ ) had a negative significant effect, while the number of people living in the household had a marginal effect ( $t = 1.85, p < .1$ ). Electricity saving was not significantly affected by whether a household is served by RPUC (that did not raise prices) or SCE (that raised prices due to the electricity crisis) which suggests that electricity saving could not be attributed to price increase. In fact, RPUC customers saved more electricity than SCE customers did, although the effect is not significant ( $t = 1.37, p = .171$ ).

Step 3 shows the effects of the moderator variables (perceptions of crisis, ownership, efficacy, and expectations for the behavior of others) as well as of the motives (utilitarianism and a fairness motive) on cooperation. The combined effect of these variables is significant, leading to a significant change in the  $F$  value ( $\Delta F = 2.80, p < .01$ ) and leading to a change in the  $R^2$  from .859 to .864. Among the variables, ownership of the problem ( $t = 1.95, p < .05$ ), efficacy beliefs ( $t = 1.67, p < .1$ ), and expectations ( $t = 1.65, p < .1$ ) had a marginal effect on electricity saving while utilitarianism had a significant, positive effect on electricity saving ( $t = 2.90, p < .01$ ), the latter providing support to Hypothesis 1. As expected, without moderation, fairness had no significant effect on cooperation ( $t = .55, p = .58$ ).

Next, the interaction effects were entered in a block, again leading to a positive significant effect on cooperation ( $\Delta F = 2.65, p < .01, \Delta R^2 = .005$ ).

Overall, the variables explained 87% of the variation ( $R^2$ ) of the electricity behavior per household in 2001.

## 7.1 Differences in interaction effects

Hypotheses 2 through 5 all state that some interaction effect between motives and situational variables will be stronger than some other interaction effect. To test this I used the statistic described in equation 10. I found all four of the hypotheses supported.

Expectations of others' cooperation influenced the cooperativeness of those with a high fairness motive more than those with a low fairness motive ( $t = 2.34, p < .05$ ) However, the effect of expectation of cooperation played no such role on the utilitarian motive ( $t = 1.17, ns$ ). As predicted in Hypothesis 4, the interaction between expected cooperation and fairness motivation ( $\beta = .112$ ) differed significantly ( $t_{diff} = 4.00, p < .001$ ) from the interaction

<i>Step</i>	<i>R</i> <sup>2</sup>	$\Delta R^2$	<i>F</i>	$\Delta F$	<i>df</i> <sub>2</sub>	<i>Variables</i>	$\beta$	<i>t</i>
1	.854		2774.65***		475			
2	.859	.005	355.18***	2.25*	468	Mean kwh 2000	-.92	-52.68***
						Cool days 2001	-.05	-2.21**
						Heat days 2001	-.00	-.18
						RPUC/SCE	.02	1.37
						Income	.02	1.20
						Own/Rent	.02	1.17
						Building type	-.02	-1.11
						# in house	-.04	-1.85 <sup>†</sup>
3	.864	.005	208.82***	2.80**	462			
						Criticality	.02	.87
						Ownership	.04	1.95*
						Efficacy	.05	1.67 <sup>†</sup>
						Expectations	.03	1.65 <sup>†</sup>
						Utilitarianism	.08	2.90**
						Fairness	.01	.55
4	.870	.006	137.64**	2.65**	454			
						Critical-Utilitarianism	.05	1.72 <sup>†</sup>
						Critical-Fair	.01	.20
						Own-Utilitarianism	.01	.33
						Own-Fair	.06	1.98*
						Efficacy-Utilitarianism	.06	2.34*
						Efficacy-Fair	.03	.73
						Expect-Utilitarianism	.03	1.17
						Expect-Fair	.11	2.61**

<sup>†</sup> $p < .1$ ; \* $p < .05$ ; \*\* $p < .01$ , \*\*\* $p < .001$

Dependent variable: negative of mean kwh billed June through August 2001

Table 6: Regression analysis of model with no-effect hypotheses

between expected cooperation and utilitarianism ( $\beta = .032$ ).

Efficacy beliefs influenced the cooperativeness of those with a high utilitarian motive more than those with a low utilitarian motive ( $t = 2.61$ ,  $p < .01$ ). However, the effect of efficacy beliefs played no such role on the utilitarian motive ( $= 0.73$ , ns). As predicted in Hypothesis 5 the interaction between efficacy beliefs and utilitarianism ( $\beta = .061$ ) differed significantly ( $t_{\text{diff}} = 3.6$ ,  $p < .001$ ) from the interaction between expected cooperation and utilitarianism ( $\beta = .025$ ).

Ownership of the problem influenced the cooperativeness of those with a high fairness motive more than those with a low fairness motive ( $t = 1.98$ ,  $p < .05$ ). However, the effect of ownership of the problem played no such role on the utilitarian motive ( $= .33$ , ns). As predicted in Hypothesis 3, the interaction between ownership of the problem and a fairness motivation ( $\beta = .055$ ) differed significantly ( $t_{\text{diff}} = 2.40$ ,  $p < .01$ ) from the interaction between ownership and utilitarianism ( $\beta = .007$ ).

Perception of criticality marginally influenced the cooperativeness of those with a high utilitarian motive more than those with a low utilitarian motive ( $t = 1.72$ ,  $p < .1$ ). However, the effect of criticality played no such role on the utilitarian motive ( $= 0.20$ , ns). As predicted in Hypothesis 2 the interaction between perception of criticality and utilitarianism ( $\beta = .052$ ) differed significantly ( $t_{\text{diff}} = 2.35$ ,  $p < .01$ ) from the interaction between expected cooperation and utilitarianism ( $\beta = .005$ ). But it must be noted the measure of perceived criticality lacks sufficient internal reliability for us to draw firm conclusions from this particular result.

## 8 Discussion and conclusions

The contribution of this paper is threefold: (1) It elaborates on the concept of fairness as a motive by distinguishing two sides of fairness: a contributory side and a detracting side. (2) It compares the effect of various situational factors on fairness and on utilitarianism in the same social dilemma context which has not been done previously; and (3) It does this in the context of a real-life, large-scale social dilemma as opposed to conducting a laboratory study that is more typically in social dilemma research. The real-life context gives external validity of the findings from laboratory studies.

By comparing utilitarianism and fairness in the same context I have shown that, although both utilitarianism and fairness can lead to cooperation in a

social dilemma, they also show different sensitivity to some aspects of a situation, leading to predictable differences in individual behavior. I found, for example, that the cooperation of fairness-motivated individuals is more positively affected by the expected behavior of others and by the ownership of a problem than the cooperation of utilitarians. I also found that the cooperation of utilitarians is more strongly affected by beliefs about the efficacy of one's contribution and by the perceived criticality of a situation than the cooperation of those with a fairness motive.

It is no surprise that situational variables moderate the relationship between utilitarianism and behavior (e.g., Dawes et al., 1977; Earley, 1994; Kramer, McClintock, and Messick, 1986), but the finding here that utilitarianism and fairness are differently sensitive to some situational variables is new. And this finding provides clear evidence that these should be taken as different motives. Previous studies tended to limit any comparison of utilitarianism with things like individualism or competitiveness (e.g., Kuhlman et al., 1992, 1986b). And very few studies treated fairness as a motive, but limited it to a constraint on other motives (e.g., de Vries and Wilke, 1992) or to a response to (un)fairness (e.g., Colquitt et al., 2001), although Van Lange (1999) is a notable exception.

Investigating utilitarianism and fairness in the same social dilemma allows us to elaborate on existing theories, such as Pruitt and Kimmel's (1977) goal/expectation theory. According to this theory, individuals are most likely to cooperate in a social dilemma if they are motivated to reach a cooperative outcome (i.e., they are utilitarians) and also expect that others will cooperate as well. Yet we have seen in this study that expectations about the behavior of others are more likely to affect the cooperative choice of those who are motivated by fairness than of those who are motivated by utilitarianism.

We can also elaborate on our understanding of the relationship between efficacy beliefs and motives. It has been widely demonstrated that large-scale social dilemmas discourage cooperation because in such situations people tend to believe that they will have little effect on reaching a cooperative outcome (e.g., Olson, 1965). Lowered efficacy beliefs were thought to decrease cooperation by moderating the relationship between motives and behavior in such a way that otherwise cooperatively-motivated individuals will reduce their effort if they do not believe that their effort will reach the desired collective outcome (Kerr, 1992). Yet, we have seen in this study that efficacy considerations do not affect the behavioral consequences of utilitarianism and fairness equally. Rather, those who are motivated by utilitarianism are more

sensitive to efficacy considerations than those who are motivated by fairness.

Knowing that both utilitarianism and fairness can lead to cooperation in a social dilemma in some contexts while differing in their behavioral consequences in other contexts may also be important for management practice. First, cooperation that emerges among fairness-motivated individuals in a social dilemma seems to be fragile in the sense that even the slightest chance that others could be perceived as not doing their fair share could lead to non-cooperation from fairness-motivated individuals (Kerr and Harris, 1996). This illustrates that even in the event of successful cooperation, it is useful to know what type of motive produced the cooperation, as some motives may lead to a more stable form of cooperation than others.

Second, different interventions are more likely to be effective in boosting cooperation when a collective is dominated by utilitarian individuals than when a collective is dominated by fairness-motivated individuals. In the former case, individuals' willingness to cooperate can be enhanced by constantly reminding people of the importance of their contributions. This approach, however, will have little impact on the cooperation of fairness-motivated individuals. In their case, the way to enhance cooperation is by making sure that they take ownership for a problem and that the cooperation of others is constantly emphasized.

## 8.1 Limitations

This study, like most, is not without limitations. A major one is that it measured individual motives using a survey technique which only allows for measuring stated preferences. Skeptics may argue that subjects knew whether they saved electricity at the time they completed the survey, so what is captured are not the motives that drove electricity saving behavior but rather a post-rationalization of the actual electricity saving behavior. Although the possibility of a post-rationalization can not be fully excluded, the fact that the measured motives lead to predicted behaviors in circumstances that are non-obvious to naive subjects suggests otherwise. Furthermore, as Baron (2001) showed, when subjects who cooperate are asked to explain their reasons for cooperating they surprisingly tend to err towards citing narrow self-interest as their motives instead of admitting altruism. In other words, a post-rationalization of a cooperative act is more likely to lead to attributing cooperation to self-interest than to utilitarianism or fairness.

A field survey of responses to a social dilemma in the wild imposes many

limits. It is just not possible to control for everything that can be controlled for in the laboratory. But the gain is real, too. I have not only provided some of the external validity called for by Pruitt and Kimmel (1977) among others, but have shown that the effects are strong enough to emerge even with a messy situation and a limited set of data collected.

A further limitation of the study is that the measure of one of the moderator variables, perceived criticality, did not reach the accepted reliability threshold and thus the findings that include this measure need to be treated with caution.

A further difficulty is that even to frame my hypotheses, I needed to present a slightly generalized view of fairness, a slightly generalized view of altruism-utilitarianism, and a slightly generalized view of public good versus resource dilemmas. While I consider each of them highly plausible and some of them well supported by existing data, I have not fully defended each of them. However, I don't consider those generalizations particularly controversial. As generalizations, they don't oppose existing claims, and for the most part they merely formalize what has been said informally. Furthermore, it is only necessary to accept some of the broad outlines of my proposed generalizations to consider this study, which does not depend on the details of those theories.

## References

- ADAMS, J. S. (1965). Inequity in social exchange. In *Advances in Experimental Social Psychology* (ed. L. Berkowitz). New York: Academic Press. Cited on: 4, 7
- ALLISON, SCOTT T. and DAVID M. MESSICK (1990). Social decision heuristics in the use of shared resources. *Journal of Behavioral Decision Making*, 3(3): 195–204. Cited on: 3
- AMBROSE, MAUREEN L. (2002). Contemporary justice research: A new look at familiar questions. *Organizational Behavior and Human Decision Processes*, 89: 803–812. Cited on: 3
- ATKINSON, JOHN W. (1982). Old and new conceptions of how expected consequences influence actions. In *Expectations and Actions: Expectancy-Value Models in Psychology* (ed. Norman T. Feather). Hillsdale, NJ: Lawrence Erlbaum Associates. Cited on: 1
- BANDURA, ALBERT (1986). *Social foundations of thought and action : A social cognitive theory*. Englewood Cliffs, N.J.: Prentice-Hall. Cited on: 14
- BARON, JONATHAN (2001). Confusion of group interest and self-interest in parochial cooperation on behalf of a group. *Journal of Conflict Resolution*, 45(3): 283–296. Cited on: 30
- BATSON, C. DANIEL and LAURA L. SHAW (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry*, 2(2): 107–122. Cited on: 9
- BORENSTEIN, SEVERIN (2001). The trouble with electricity markets and California’s electricity restructuring disaster. Working Paper WP-081, University of California Energy Institute, Berkeley, CA. (<http://www.ucei.org/PDF/pwp081r1.pdf>). Cited on: 17
- CHATMAN, JENNIFER A. and SIGAL G. BARSAD (1995). Personality, organizational culture, and cooperation: Evidence from a business simulation. *Administrative Science Quarterly*, 40(3): 423–443. Cited on: 12

- COHEN-CHARASH, YOCHI and PAUL E. SPECTOR (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 86(2): 278–32. Cited on: 3
- COLQUITT, JASON A., DONALD E. CONLON, MICHAEL J. WESSON, CHRISTOPHER O. L. H. PORTER, and K. YEE NG (2001). Justice at the millenium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86(3): 425–445. Cited on: 2, 29
- CROSON, RACHEL (1998). Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. Working paper, University of Pennsylvania. Cited on: 14
- DAWES, ROBYN M., JEANNE MCTAVISH, and HARRIET SHAKLEE (1977). Behavior, communication, and assumptions about other people’s behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, 35(1): 1–11. Cited on: 14, 29
- DAWES, ROBYN M., JOHN M. ORBELL, RANDY T. SIMMONS, and ALPHONS J. C. VAN DE KRAGT (1986). Organizing groups for collective action. *The American Political Science Review*, 80(4): 1171–1185. Cited on: 15
- DEUTSCH, MORTON (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal Social Issues*, 31(3): 137–149. Cited on: 4
- DEUTSCH, MORTON (1985). *Distributive justice: A social psychological perspective*. New Haven, CT: Yale University Press. Cited on: 4
- EARLEY, CHRISTOPHER P. (1994). Self or group? Cultural effects of training on self-efficacy and performance. *Administrative Science Quarterly*, 39(1): 89–117. Cited on: 29
- ELSTER, JON (1989). *The Cement of Society: A study of social order*. Studies in Rationality and Social Change. Cambridge: Cambridge University Press. Cited on: 3
- GREENBERG, JERALD (1978). Effects of reward value and retaliative power on allocation decisions: Justice, generosity, or greed? *Journal of Personality and Social Psychology*, 36(4): 367–379. Cited on: 4

- HARDIN, G. (1968). The tragedy of the commons. *Science*, 162: 1243–1248.  
Cited on: 1
- HEGTVEDT, KAREN A. and KAREN S. COOK (2001). Distributive justice: Recent theoretical developments and application. In *Handbook of justice research in law* (eds. Joseph Sanders and V. Lee Hamilton), pp. 93–132. New York: Kluwer Academic. Cited on: 4
- HOFFMAN, MARTIN L. (1981). Is altruism part of human nature? *Journal of Personality and Social Psychology*, 40(1): 121–137. Cited on: 9
- JOHNSON, AARON C., MARVIN B. JOHNSON, and RUEBEN C. BUSE (1987). *Econometrics: Basic and Applied*. New York: MacMillan. Cited on: 24
- KERR, NORBERT and SUSAN E. HARRIS (1996). Why do cooperators cooperate? Efficacy as a moderator of social motives effects. In *Frontiers in Social Dilemmas Research* (eds. Wim Liebrand and David Messick), pp. 101–115. Berlin: Springer. Cited on: 15, 30
- KERR, NORBERT L. (1992). Efficacy as a causal and moderating variable in social dilemmas. In *Social Dilemmas: Theoretical issues and research findings* (eds. Wim B. G. Liebrand, David M. Messick, and Henk A. M. Wilke), International Series in Experimental and Social Psychology, pp. 59–80. Oxford: Pergamon Press. (Selected from the Third conference on Social Dilemmas (1991), Groningen). Cited on: 14, 15, 29
- KRAMER, RODERICK M., MARILYNN B. BREWER, and BENJAMIN A. HANNA (1996). Collective trust and collective action. The decision to trust as a social decision. In *Trust in organizations. Frontiers of theory and research* (eds. Roderick M. Kramer and Tom R. Tyler), pp. 357–389. Thousand Oaks: Sage. Cited on: 1
- KRAMER, RODERICK M., CHARLES G. McCLINTOCK, and DAVID M. MESSICK (1986). Social values and cooperative response to a simulated resource conservation crisis. *Journal of Personality*, 54(3): 576–592. Cited on: 29
- KUHLMAN, D. M., C. R. CAMAC, and D. A. CUNHA (1986a). Individual differences in social orientation. In *Experimental Social Dilemmas* (eds.

- Henk A. M. Wilke, Dave M. Messick, and Christel G. Rutte). Frankfurt am Main: Verlag Peter Lang. Cited on: 9
- KUHLMAN, D. MICHAEL, CLIFFORD BROWN, and PAUL TETA (1992). Judgments of cooperation and defection in social dilemmas: The moderating role of judge's social orientation. In *Frontiers in Social Dilemmas Research* (eds. Wim B. G. Liebrand, David M. Messick, and Henk A. M. Wilke), pp. 111–132. Oxford: Pergamon Press. (Selected from the Third conference on Social Dilemmas (1991), Groningen). Cited on: 2, 9, 11, 29
- KUHLMAN, D. M[ICHAEL]., C. R. CAMAC, and D. A. CUNHA (1986b). Individual differences in social orientation. In *Experimental Social Dilemmas* (eds. Henk A. M. Wilke, Dave M. Messick, and Christal G. Rutte), volume 3 of *Psychologie des Entscheidungsverhaltens und des Konfliktes*, pp. 151–176. Frankfurt am Main: Peter Lang. (Proceedings of the 1984 Conference on Social Dilemmas held at the University of Gronigen). Cited on: 2, 9, 10, 11, 29
- KUHLMAN, D. MICHAEL and ALFRED F. J. MARSHELLO (1975). Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma. *Journal of Personality and Social Psychology*, 32(5): 922–931. Cited on: 9
- LERNER, MELVIN (1977). Some hypotheses as to its origins and forms. *Journal of Personality*, 45(1): 1–52. Cited on: 3
- LIEBRAND, WIM B. (1984). The effect of social motives, communication and group size on behaviour in an n-person multi-stage mixed motive game. *European Journal of Social Psychology*, 14: 239–264. Cited on: 9, 10, 11
- LIND, E. ALLAN and TOM R. TYLER (1988). *The social psychology of procedural justice*. New York: Plenum Press. Cited on: 3
- MACCRIMMON, KENNETH R. and DAVID M. MESSICK (1976). A framework for social motives. *Behavioral Science*, 21: 86–100. Cited on: 9, 10
- MESSÉ, L. A. and J. M. SIVACEK (1979). Predictions of others' responses in a mixed-motive game: Self justification or false consensus. *Journal of Personality and Social Psychology*, 37: 602–607. Cited on: 14

- MESSICK, DAVID M. (1980). Equality as a decision heuristic. In *Psychological Perspectives on Justice* (eds. B. A. Mellers and J. Baron), pp. 11–31. New York: Cambridge University Press. Cited on: 4
- MESSICK, DAVID M. and MARILYNN B. BREWER (1983). Solving social dilemmas: Individual and collective approaches. In *Review of Personality and Social Psychology* (ed. L. Wheeler). Beverly Hills, CA: Sage publications. Cited on: 1
- MESSICK, DAVID M. and KAREN S. COOK (eds.) (1983). *Equity theory : Psychological and sociological perspectives*. New York: Praeger. Cited on: 4
- MESSICK, DAVID M. and CHARLES G. MCCLINTOCK (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology*, 4(1): 1–25. Cited on: 1
- MESSICK, DAVID M. and KEITH P. SENTIS (1979). Fairness and preference. *Journal of Experimental Social Psychology*, 15(4): 418–434. Cited on: 3, 7
- MILES, EDWARD W., JOHN D. HATFIELD, and RICHARD C. HUSEMAN (1989). The equity sensitivity construct: Potential implications for worker performance. *Journal of Management*, 15(4): 581–588. Cited on: 3
- OLSON, M. (1965). *The Logic of Collective Action*. Cambridge, MA: Harvard University Press. Cited on: 15, 29
- ORGAN, DENNIS W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books. Cited on: 11
- PRUITT, DEAN G. and MELVIN J. KIMMEL (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28: 363–392. Cited on: 1, 13, 29, 31
- REEVE, HUDSON KERN (1998). Acting for the good of others: Kinship and reciprocity with some new twists. In *Handbook of evolutionary psychology: Ideas, issues, and applications* (eds. Charles B. Crawford and Dennis L. Krebs), pp. 43–85. Mahwah, NJ: Lawrence Erlbaum Associates. Cited on: 9

- ROSS, LEE and RICHARD E. NISBETT (1991). *The person and the situation. Perspectives of Social Psychology*. New York: Mc-Graw Hill. Cited on: 12
- SCHROEDER, DAVID A., MARK E. SIBICKY, and MARY ELIZABETH IRIS (1995). A framework for understanding decisions in social dilemmas. In *Social Dilemmas: Perspectives on Individuals and Groups* (ed. David A. Schroeder), pp. 183–199. Westport, CT: Praeger. Cited on: 13
- SCHWARTZ, SHALOM H. and AVI GOTTLIEB (1980). Bystander anonymity and reaction to emergencies. *Journal of Personality and Social Psychology*, 39: 418–430. Cited on: 13
- SIMON, HERBERT A. (1990). A mechanism for social selection and successful altruism. *Science*, 250: 1665–1668. Cited on: 9
- SIMON, HERBERT A. (1993). Altruism and economics. *American Economic Review*, 83(2): 156–161. Cited on: 9
- SMITH, VINCENT, MICHAEL KEHOE, and MARY CREMER (1995). The private provision of public goods: Altruism and voluntary giving. *Journal of Public Economics*, 58: 107–126. Cited on: 11
- UNGER, LYNETTE S. (1991). Altruism as a motivation to volunteer. *Journal of Economic Psychology*, 12(1): 71–100. Cited on: 11
- VAN LANGE, PAUL 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(2): 337–349. Cited on: 2, 5, 6, 9, 29
- VAN VUGT, MARK, REE M. MEERTENS, and PAUL A. M. VAN LANGE (1994). Commuting by car or by public transportation? (a)n interdependence theoretical approach. In *Social Dilemmas and Cooperation* (eds. Ulrich Schulz, Wulf Albers, and Ulrich Mueller), pp. 291–309. Berlin: Springer-Verlag. Cited on: 13
- DE VRIES, SJERP and HENK A. M. WILKE (1992). Constrained egoism and resource management under uncertainty. In *Social dilemmas: Theoretical issues and research findings* (eds. Wim B. G. Liebrand and David M. Messick). Elmsford, NY: Pergamon Press. Cited on: 29