Article

Do I Really Want to Know? A Cognitive Dissonance-Based Explanation of Other-Regarding Behavior

Astrid Matthey and Tobias Regner [*]

Max-Planck-Institute of Economics, Kahlaische Str. 10, 07745 Jena, Germany

[*] Author to whom correspondence should be addressed; E-Mail: regner@econ.mpg.de;
Tel.: 03641/686-635; Fax: 03641/686-667.

Abstract: We investigate to what extent genuine social preferences can explain observed other-regarding behavior. In a dictator game variant subjects can choose whether to learn about the consequences of their choice for the receiver. We find that a majority of subjects showing other-regarding behavior when the payoffs of the receiver are known, choose to ignore these consequences if possible. This behavior is inconsistent with preferences about outcomes. Other-regarding behavior may also be explained by avoiding cognitive dissonance as in Konow (2000). Our experiment's choice data is in line with this approach. In addition, we successfully relate individual behavior to proxies for cognitive dissonance.

Keywords: social preferences; other-regarding behavior; experiments; cognitive dissonance

1. Introduction

By now a substantial body of literature exists on other-regarding behavior. While the empirical evidence is abundant and affirmative, the theoretical side remains less unified.[1] The aim of this paper is to further contribute to the research on social preferences. In particular, we try to shed some more light on the actual process that leads to behavior that is perceived as "other-regarding".

---

[1] Two approaches can be distinguished, outcome-based and intentions-based models. Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) are seminal papers of the first strand, Rabin (1993), Levine (1998), Dufwenberg and Kirchsteiger (2004), and Falk and Fischbacher (2006) of the second. Both approaches have their respective advantages and drawbacks as illustrated in surveys on the literature (Camerer, 2003; or Fehr and Schmidt, 2003).

Participants in our experiment face some standard "full information" allocation situations, that is, situations where they are fully informed about the payoff of the participant affected by their choice. In addition, they face allocation situations where they have the possibility to remain uninformed about the consequences on the other participant's payoff. People know that their behavior affects others, but they can ignore the exact consequences of this behavior. They can—but do not have to—find out which of several possible effects their behavior has on others.

Several previous studies analyzed subjects' behavior when a costly exit option to get out of a dictator game is provided (Dana *et al.*, 2006; Lazear *et al.*, 2006; Broberg *et al.*, 2007). Our design differs from such 'dictator exit' situations. We are rather interested in whether there is a change in subjects' behavior given that they remain in an allocation situation: Do people who show other-regarding behavior under "full information" choose to find out the consequences of their choice if they can stay uninformed? In other words, are they "genuine pro-socials"? Or do they act pro-socially under full information, yet prefer to remain ignorant when given the chance (taking negative consequences for others into account)? Such 'strategic ignorance' has been studied by Dana *et al.* (2007) and Larson and Capra (2009) in a between-subjects design.

Our goal is to shed light on the driving motivational forces that distinguish these "ignorers" from "genuine pro-socials" and we relate our findings to cognitive dissonance, a psychological theory introduced by Leon Festinger (1957). A person experiences cognitive dissonance when she holds two psychologically conflicting cognitions, say, she finds a certain task boring, but claims that it was interesting as an internal justification for actually doing it (Festinger, 1957). In its modern version of the theory (Aronson, 1994; Beauvois and Joule, 1996)[2] cognitive dissonance primarily revolves around the self and a piece of behavior that violates that self-concept. This modern view is most related to our context. It would interpret other-regarding behavior as being driven by individuals' desire to avoid a divergence between the behavior they consider appropriate and the behavior they actually choose, since such a divergence would lead to the unpleasant feeling of cognitive dissonance.

In our experiment cognitive dissonance may affect subjects' behavior in the following way. Some subjects prefer a fair split under full information but choose to ignore if given the chance. Then they will have two inconsistent psychological cognitions in the second situation: their norm to behave fairly, and their choice deviating from that norm. Hence, they should experience cognitive dissonance, a psychological cost. Ignoring pro-socials bear this cost as it is outweighed by the higher material payoff. while for genuine pro-socials the potential psychological conflict is too high and they choose a fair outcome also if they could have remained ignorant.

The results of the experiment show that indeed a significant number of subjects who show other-regarding behavior under full information prefer to remain ignorant if this is possible. Cognitive dissonance can explain such a split into ignoring and genuine pro-socials, while models based on preferences about outcomes cannot. In these models, a subject's preference for a certain outcome distribution does not depend on whether the information regarding these outcomes is provided

---

[2]See Harmon-Jones and Mills (1999) for a review of the current state of dissonance theory. In contrast to social psychology where it has been frequently applied only few articles in economics use it to explain decision making (see for instance Akerlof and Dickens, 1982; Rabin, 1993; Konow, 2000; and Oxoby, 2003, 2004)

immediately or can be obtained at no cost. Hence, the opportunity to ignore cannot explain a difference in choices.

Our results are in line with Dana *et al*. (2007) where subjects also have the opportunity to remain ignorant about the exact consequences of their behavior for their partner. Similar to us, the authors find that this possibility reduces the extent of other-regarding behavior. However, since the authors use a between-subjects design and assess only subjects' choices, their data does not allow a test of the theory. Konow (2000) provides a theoretical model within the cognitive dissonance framework and offers supporting experimental evidence based on choice data. We also find that cognitive dissonance is useful to organize our choice data of ignoring behavior. In addition we successfully relate this behavior to proxies of cognitive dissonance.

Our moral dilemma experiment reflects features of many decisions we make every day. We often know that a choice has implications for others, but can remain uninformed about the exact consequences of this choice. Whenever we shop for food, clothes, *etc*., we face the decision of whether to pay higher prices for goods that were produced under higher social and/or ecological standards. This means that with each purchase decision we choose whether to share our money with others, who then get higher wages, better social insurance, health care *etc*. Before we make this decision, however, we decide whether to pay attention to the way of production, whether a product is traded fairly, or whether the fair trade label a product bears is credible. We can avoid the sharing decision if we simply pick the product which we like best, and ignore the way it was produced and traded. This is different from a situation where people have an alternative where their behavior does not have a direct impact on others, like the beggar on the street, whom we can avoid by crossing the road. This costless alternative is rarely available when we purchase food, clothes or other consumption goods.

The paper is organized as follows. In Section 2 we describe the experimental design. Section 3 relates the experiment to the theoretical background and derives behavioral predictions. Results are presented in Section 4 and discussed in Section 5. Section 6 concludes.

2. Experiment

2.1. Participants and Procedures

90 participants were recruited among students from various disciplines at the local university. In each session gender composition was approximately balanced and subjects took part only in one session. Participants were recruited using the ORSEE software (Greiner, 2004). The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007) and took, on average, 50 minutes. The average earnings in the experiment have been €7.61 or $10.65. The show-up fee for the experiment was €2.50 or $3.50.

Upon arrival at the laboratory subjects were randomly assigned to one of the computer terminals. Each computer terminal is in a cubicle that does not allow communication or visual interaction among the participants. Participants were given time to privately read the instructions. Subjects were allowed to privately ask for clarifications about the instructions. In order to check the understanding of the instructions subjects were asked to answer some control questions. After all subjects had answered the questions correctly the experiment started.

At the end of the experiment subjects were paid in cash according to their performance. Privacy was warranted during the payment phase.

2.2. Design

The experiment consisted of four periods. In periods 1 to 3, each subject chose among five options, A to E. Each option assigned one outcome to the subject herself, and one to another person, named participant Y. In circular matching fashion a third participant, called participant X, also assigned one outcome to himself and one to the subject. This circular matching was introduced to avoid possible feelings of reciprocity that could result if subjects were matched in pairs.[3] Participants X and Y were chosen randomly, and unknown to the subject. Subjects were matched using a perfect stranger design.

In each period (1-3), the five outcomes assigned by the options were 1 EC (experimental currency), 1 EC, 4 EC, 4 EC and 7 EC, for all participants. This was common knowledge. Each option a subject could choose assigned one of these outcomes to the subject herself, and one to the subject in the role of participant Y. How the two sets of outcomes were combined was determined randomly. For example, the option paying 7 EC to the subject herself could induce an outcome of 7 EC for Y (with 20% chance), of 4 EC (40% chance) or of 1 EC (40% chance). These probabilities were also stated in the instructions (see appendix; the instructions also contain tables that illustrate the available choices for the subjects).

In *periods 1 and 3*, subjects could initially only see their own outcomes, but not the outcomes their choice would assign to Y. Subjects could then decide whether to i) choose an option immediately, *i.e.*, without knowing the consequences for Y, ii) uncover the outcomes that each option assigned to participant Y, and then choose an option, or iii) see the outcome pairs from which participant X could choose, and then choose an option (their own, not that of X). Subjects were aware that knowing the options of X would not allow them to influence his choice. Hence, uncovering the outcome pairs of X was rather an excuse for not uncovering the consequences of one's choice for Y.

In *period 2*, the outcomes of the subject herself and participant Y were visible immediately, and subjects directly chose an option. In fact, the options (outcome pairs) in periods one and two were the same. Only their order was reversed to make this fact less obvious.

In *period 4* subjects faced a series of options. First they had to choose between an outcome of 10 EC for themselves and 0 for participant Y, versus equal outcomes for both. The equal outcomes increased from 0 EC for both to 10 EC for both in steps of 0.5 EC. Second, subjects had to choose between the following pairs of outcomes, where the first number of each pair denotes the subject's own outcome (in EC) and the second the outcome of participant Y: (4,4) *vs.* (4,7); (4,1) *vs.* (4,7); (7,1) *vs.* (4,4); (7,1) *vs.* (4,7). Only one of the 25 choices in period 4 was paid. We used the strategy method in order to get a more comprehensive picture of the subjects' preferences.

After all choices were made, one period was chosen randomly for each individual. From this period, the individual received either the outcome she assigned to herself, or the outcome that participant X assigned to her. If period 4 was drawn, the decision to be paid was determined randomly in a second step.

---

[3]This means that in each period, each subject received one outcome from a participant called X, and chose an option which assigned one outcome to herself and one to another participant called Y. Participants X and Y were unrelated (the matching groups contained 18 subjects).

At the end of the experiment subjects filled in a questionnaire. It first assessed subjects' emotions in the decision situation for the items *happiness, anger, guilt, pride, irritation, shame, determination, confusion* (scales from 1 to 5). The relevant items for our study are *negative emotions towards others* (anger, irritation; Cronbach's $\alpha = 0.79$), *negative self-related emotions* (shame, guilt; $\alpha = 0.69$) and *positive self-related emotions* (pride, happiness; as α is not sufficiently high (0.49) we only use one of the items in the following analysis. Both produce qualitatively equal results.). Next, subjects were asked *how hard it was for them to decide* whether they wanted information on the outcomes of participant Y, or on the options of participant X.[4] Then, we asked—among other things—for the *share of subjects they expected to have uncovered* the outcomes of participant Y in periods 1 and 3, how important *principles in life* are to them, and how relevant the *opinion of others* is to them. Finally, we asked for some background information (field of studies, gender). See also appendix B for the full text of the post-experimental questionnaire.

3. Theoretical Background

3.1. The Cognitive Dissonance-Based Model of Konow (2000)

The model of Konow (2000) uses the concept of cognitive dissonance to explicitly explain behavior in dictator games. It is therefore well suited as a benchmark against which to compare our data. In Konow's model, individuals maximize a utility function that is increasing in the share of money they keep for themselves, y, and decreasing in the amount of cognitive dissonance they experience and the extent of self-deception they engage in. People experience cognitive dissonance if the share y an individual keeps for herself differs from the share she believes is fair to keep, ϕ. An individual engages in self-deception if she makes herself believe that a certain share ϕ is fair to keep, but this share differs from the fair entitlement η. An important assumption of the model is the general definition of this fair entitlement, for all individuals and across all situations, by the *accountability principle*: individuals' payoffs should not depend on variables they have no control over.

In addition to the share y and the belief ϕ, an individual's dissonance function f depends on the parameter α which specifies her sensitivity to cognitive dissonance. Since α may vary across individuals and contexts, so does the amount of cognitive dissonance an individual experiences given y and ϕ. Similarly, given the fair entitlement η and the individual's belief about the fair share ϕ, the (emotional) costs of self-deception that arise if $\eta < \phi$ depend on the parameter β. The individual then solves the problem

$$\text{Max}_{y,\phi} \quad u(y, \phi, \eta, \alpha, \beta) \equiv v(y) - f(y - \phi, \alpha) - c(\phi - \eta, \beta) \tag{1}$$

that is, she chooses the share to keep and her belief of the fair share. According to these choices she obtains utility from the monetary payoff, $v(y)$, and incurs potential costs of cognitive dissonance, $f(y - \phi, \alpha)$, and self-deception, $c(\phi - \eta, \beta)$.

---

[4]The precise (translated) text of this question is: "Did you find it easy to decide in these situations [period 1 or 3], whether you want to receive information about the payoff of participant Y or the payoff table of participant X?"

3.2. Predictions for The Experiment

What does Konow's model predict for the behavior in the experiment? Consider what happens in the ignorance (period 1 and 3) compared to the open situation (period 2 and 4). First, according to the accountability principle, the entitlement η does not change, since the other player is not responsible for the option to ignore. Further, the sensitivity to cognitive dissonance, α, cannot be expected to change either, since an additional option will not change the general discomfort involved in choosing a share that is considered unfair. No matter what the circumstances are, once a subject considers ϕ the fair share but keeps $y > \phi$, this should induce the same unpleasant feelings irrespectively of how η and ϕ evolved. However, subjects may perceive different shares ϕ as fair in the two situations. In particular, the option to ignore the other's payoff may reduce the costs of self-deception, β, and lead to a higher optimal ϕ in periods 1 and 3 than in periods 2 and 4. If this is the case, *i.e.*, β decreases at least marginally, subjects who choose to behave fairly in the open situation may find it optimal to ignore if given the chance, due to the lower costs of self-deception.

We would then observe two classes of pro-social subjects: For the first class of subjects, if they had made a selfish choice in the open situation, the disutility from cognitive dissonance and self-deception would outweigh the monetary utility *by far*. For them, even if β decreases in the ignorance situation, the balance is still strongly in favor of a pro-social (non-ignoring) choice. In contrast, if the second class of pro-socials had made the selfish choice in the open situation, the disutility from cognitive dissonance and self-deception would only *slightly* outweigh the monetary utility. For these subjects, if β decreases in the ignorance situation, the balance may tip towards an ignoring choice. Although not all subjects of this class may finally decide to ignore, the decision is much harder for them than for the first class of pro-socials because the utility expected from the monetary payoff and the (dis-)utility expected from cognitive dissonance and self-deception are of similar size. Since all subjects who ignore in situations 1 or 3 come from this class, but only some who do not ignore, on average ignoring pro-socials have more difficulty making their decisions.

Further, as long as β does not decrease to zero, ignoring pro-socials experience cognitive dissonance from their choice to ignore, as this choice is not in line with the behavior determined by the accountability principle. Since cognitive dissonance is a "negative drive state" (Festinger, 1957), this experience should result in unpleasant feelings for ignoring pro-socials.

- Some subjects behaving pro-socially in the open situations (periods 2 and 4) ignore in period 1 or 3. For these ignoring pro-socials, decisions are more difficult than for genuine pro-socials. They also experience higher cognitive dissonance on average.

To test the first part of the hypothesis, we check whether subjects behaving pro-socially in the open situations always choose to uncover in periods 1 and 3, or whether some ignore if given the chance. For the second part, we analyze whether these ignoring pro-social subjects take longer to decide upon their choice and report more difficulty making this choice than the pro-socials who do not ignore. Finally, we test whether the negative drive state ignoring pro-socials should be in can be detected in their emotions. To do so we compare the (positive and negative) self-related emotions[5] that ignoring pro-socials report

---

[5]We do not predict an effect on the negative emotion *towards others*, since it is not clear how our classification of subjects is related to the emotions that arise if a subject receives a (possibly low) payoff due to the decision of another player.

in the post-experimental questionnaire (whether they were proud of themselves/happy, or whether they felt ashamed/guilty) to those reported by genuine pro-socials.

Subjects who behave selfishly in the open situation experience cognitive dissonance if $\alpha > 0$ and $\beta > 0$, that is, if they are sensitive to cognitive dissonance and unable to deceive themselves at no cost. For these subjects, ignorance is the dominant choice in periods 2 and 4 if β decreases at least marginally, *i.e.*, if self-deception becomes less costly under ignorance. In contrast, if $\alpha = 0$, $\beta = 0$, or β does not decrease under ignorance, selfish subjects are indifferent between ignoring and not ignoring, since both choices result in the same utility. This means that for both kinds of selfish subjects choices in periods 1 and 3 are simple as they either face a dominant choice or a choice between equal options. However, except for subjects with $\alpha = 0$ or $\beta = 0$, selfish subjects of both types experience cognitive dissonance, since their choices deviated from the choices implied by the accountability principle. Accordingly, they should experience unpleasant emotions as a result of their choice.

Some subjects behaving selfishly in the open situations (periods 2 and 4) ignore when possible (periods 1 and 3). For all selfish subjects, decisions are as easy as for genuine pro-socials. However, selfish subjects tend to experience higher cognitive dissonance than genuine pro-socials, independently of the choice to ignore.

To test the first part of the hypothesis, we analyze whether subjects behaving selfishly in the open situations choose to ignore if given the chance. For the second part, we analyze whether selfish subjects—independently of their choice to ignore—decide upon their choice as quickly and report as little difficulty making this choice as do genuine pro-socials. Finally, we test whether the negative drive state pro-self subjects should be in due to their deviation from the accountability principle can be detected in their emotions. Again, we compare the (positive and negative) self-related emotions that pro-selfs report in the post-experimental questionnaire to those reported by genuine pro-socials.

Other variables from the post-experimental questionnaire allow us to make additional behavioral predictions that are not directly implied by Konow's model. If social norms positively affect behavior, the share of others that a subject expects to ignore will (weakly) increase her own probability to ignore. Hence, subjects who ignore will expect a (weakly) higher share of other ignorers.

Further, if the accountability principle holds, ignoring pro-socials and selfish subjects are expected to attach lower relevance to personal principles and consequently make choices that are less in line with these principles. Genuine pro-socials are expected to attach more weight to these principles, leading to choices more in line with them.

Finally, it could play a role how much attention subjects pay to others' opinions. In this case, if the accountability principle is generally accepted/applied, a positive relation between pro-social behavior and others' opinion would be expected, since these opinions should reflect the accountability principle.

4. Results

Generally, our interest is whether we can in fact observe "ignoring" behavior among individuals. In a first step we use data from period 4 of the experiment in order to classify the 90 subjects into selfish and

pro-social types. This classification is based on subjects' preferences for a (7,1) payoff or an equal split of (4,4) elicited in period 4.[6] It produces 54 pro-self and 36 pro-social types.

Next, we are interested in the behavior of these pro-social types, when they face a situation as in period 1 or 3 (consequences of choices for the other participant are unknown, yet a mouse click will reveal them). If a pro-social participant does not uncover his/her own payoff table in either period 1 or 3, then we classify him/her as an *ignoring pro-social*. *Genuine pro-socials* always uncover the additional information so they can consider the other participant's payoff in their decision. We find 23 ignoring pro-socials and 13 genuine pro-socials.

Selfish individuals may also be different in the way they behave in periods 1 and 3. Uncovering the payoff table means getting to know the payoffs of participant Y. This bears no importance for those selfish individuals who are ex-ante convinced that they will choose the option containing 7 for themselves no matter what. It may make a difference for individuals who are generally selfish, but more liable to be affected by the context. They would prefer not to uncover in order to avoid knowledge about the context, *i.e.* the consequence for the other participant. Hence, if a selfish participant does not uncover in periods 1 and 3, we classify him/her as an *unstable pro-self* (31 participants) and if (s)he uncovers at least once we classify him/her as a *convinced pro-selfs* (23 participants). See Table 1 for an overview of the type categorization.

Table 1. Categorization of the 90 subjects.

| | | In period 4 preference for | |
		(7,1) over (4,4)	(4,4) over (7,1)
In periods 1/3	preference to ignore	31 unstable pro-selfs	23 ignoring pro-socials
	preference to uncover	23 convinced pro-selfs	13 genuine pro-socials

The categorization into pro-social and pro-self types is based on subjects' binary choice between (7,1) and (4,4) in period 4. How consistent is this with respect to subjects' choices in the previous periods? If a subject uncovered in period 1 92.9% of pro-self and 94.7% of pro-social individuals acted consistent with their type. These subjects saw the equivalent options in the payoff table in period 2. 92.9% of pro-self and 94.1% of pro-social individuals made the same choice. If a subject did not uncover in period 1, we can check whether the choice in period 2 is consistent with the categorization. 93.8% of pro-socials and 81.6% of the pro-selfs are consistent. Finally, 89.5% pro-social and all pro-self subjects who uncovered in period 3 are consistent with their type.

With the participants classified into the types convinced/unstable pro-selfs, and ignoring/genuine pro-socials, we proceed to analyze which variables distinguish between these types. We use data from the experiment and answers in the post-experimental questionnaire as independent variables. As proxies

---

[6]Alternatively, the strategy method-induced preferences for the decision between a (10,0) split and an equal split of (x,x), where x increased in steps of €0.5 could be used. A reasonable cutoff level would be €6.5 to distinguish between selfish (always choice of (10,0) if $x \leq 6.5$) and pro-social (always choice of the equal split if $x > 6.5$) participants. The following results are robust with respect to this alternative categorization.

for cognitive dissonance[7] we use subjects' experienced *conflict* (their answer to the question how hard it was for them to decide whether they wanted information on the outcomes of participant Y, or on the options of participant X) and the computed *reaction times* of participants.[8] These are the times a participant spent in front of the first screen (possible decisions: uncover payoffs of participant Y, uncover the decision table of participant X or choose an option (A to F) straightaway) and in front of the second screen (possible decisions: choice of A to F with the payoffs of Y uncovered or the decision table of X shown). Average reaction times for the types are 64.39 (convinced pro-selfs), 80.32 (unstable pro-selfs), 86.04 (ignoring pro-socials), 45.84 (genuine pro-socials). Reaction times of genuine pro-socials are significantly different from those of unstable pro-selfs (ranksum test, $p = .01$) and ignoring pro-socials ($p = .01$), but not compared to convinced pro-selfs ($p = .12$). There are no significant differences in the reaction times of convinced pro-selfs, unstable pro-selfs and ignoring pro-socials. The distributions, see Figure 1, are slightly right-skewed. Hence, we use the logarithm of the reaction times in the further analysis.

Figure 1. Distribution of reaction times by types.
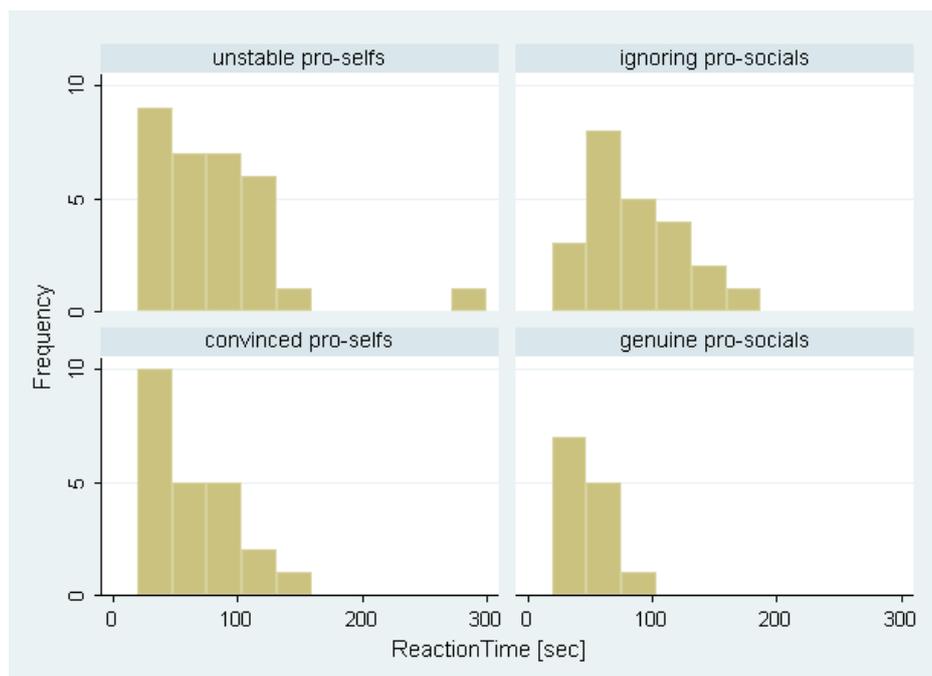
Moreover, we use positive/negative self-related emotions, negative emotions towards others, the *share of subjects they expected to have uncovered* the outcomes of participant Y in periods 1 and 3, how important *principles* are to them, and how much attention they pay to *others' opinions*.

We run a multinomial logit regression in order to distinguish between individuals who clearly prefer to uncover (genuine pro-socials), and those who prefer not to uncover (ignoring pro-socials and unstable pro-selfs as they would feel obliged to share or at least feel somewhat bad about not sharing), and those who do not care (convinced pro-selfs). The base outcome of the multinomial logit regression is the type

---

[7]To be more precise, for the tradeoff between monetary utility and disutility from cognitive dissonance and self-deception
[8]Rubinstein (2007) propagates the use of response times in order to explore the deliberation process of decision makers. A positive correlation between the experienced conflict and response time has also been found by Branas-Garza *et al*. (2007).

genuine pro-socials, *i.e.* individuals who are somewhat sure what they do/ought to do in such a situation. See Table 2 for the results.

Table 2. Multinomial logit with type as dependent variable.

	ignoring pro-socials		unstable pro-selfs		convinced pro-selfs	
	coeff.	st.error	coeff.	st.error	coeff.	st.error
ExperiencedConflict	1.895	.7261 ***	.3077	.8929	.9842	.6976
log(ReactionTime)	3.762	1.197 ***	2.355	1.166 **	1.278	.9808
ShareIgnorers	2.267	3.239	8.363	2.934 ***	3.541	2.781
NegEmoOther	−1.021	1.077	−.0767	1.147	−.211	1.125
NegEmoSelf	.2029	1.067	.9026	1.304	1.119	1.114
PosEmoSelf	−1.431	.5579 ***	−1.496	.726 **	−1.515	.5656 ***
Principles	.5395	.7959	−.2346	.8143	−.7227	.575
OtherOpinion	.481	.4964	.5386	.5067	.1189	.4432
studies business/econ	1.974	1.353	3.431	1.397 **	2.872	1.259 **
female	−.403	1.05	−1.092	1.076	−.6579	.8694
constant	−19.46	7.094 ***	−12.31	7.086 *	−3.531	4.837

Pseudo R2 = 0.334	Log pseudolikelihood = −80.54
$N = 90$	significance levels: $*** = 1\%, ** = 5\%, * = 10\%$

The regression shows a positive and significant effect of the conflict variable and the reaction time (both at the 1%-level) for ignoring pro-socials. It appears that participants classified as ignoring pro-socials experience a conflict. They also tend to require more time when they face the period 1/3 decisions (whether to uncover or select one of the options right away). These results are in line with the prediction of the cognitive dissonance-based model (hypothesis 3.2). Some participants who otherwise act socially-minded do not uncover (our ignoring pro-socials type). They should experience a high level of cognitive dissonance due to the two conflicting psychological cognitions: A general sense for sharing equally that was however dismissed in a situation where ignoring the consequences for the other was facilitated and a material advantage was gained at the expense of the other. Such a dissonance between these two inconsistent cognitions should result in experiencing a conflict and in longer times to decide between the options. Our experimental results confirm this as well as the resulting negative drive state that should have implications on subjects' self-related emotions. Ignoring pro-socials are significantly less proud/happy (1%-level) than genuine pro-socials, although there is no difference in negative self-related emotions.

Indications for such an experienced dissonance can also be found for the pro-self types. They report significantly less pride/happiness about their decision than genuine pro-socials (again there is no difference in negative self-related emotions). As well in line with hypothesis 3.2 pro-selfs do not seem to express that they had difficulties making the decisions to uncover the payoff table. While unstable

pro-selfs need significantly more time than the comparison group of genuine pro-socials (5%-level), convinced pro-selfs do not need significantly more time than the genuine pro-socials.

Participants' beliefs about the share of others who ignore is significant at the 1%-level for the unstable pro-self type. However, the data does not support such a correlation for ignoring pro-socials. None of the negative emotions experienced appear to distinguish between the types. Also participants' personal principles and how much attention they pay to others' opinions do not seem to matter. Results are robust to leaving out the control variables (gender, business/economics student). Also an alternative specification that treats ignoring pro-socials and convinced pro-selfs as the joint base outcome type of the regression delivers qualitatively comparable results.

5. Discussion

This section discusses our results in light of the existing related literature. We compare the behavioral pattern found in our experiment to the previously mentioned studies of moral dilemmas, and we discuss alternative explanations for 'dictator exit' / 'strategic ignorance' brought forward by this literature. Finally, we analyze to what extent the cognitive dissonance-based model of Konow (2000) is able to explain our data.

5.1. Previous Empirical Evidence

Our results are generally consistent with findings of the previously mentioned studies. In particular our categorization of subjects into types delivers results that are in line with Dana *et al.* (2007), the other study of 'strategic ignorance'.[9] In their baseline treatment, a transparent situation, about a quarter of subjects (5 out of 19) behave self-interestedly, and about a third of subjects behave in a pro-social way in their 'strategic ignorance' treatments. In our within-subjects design 23 out of 90 subjects are categorized as convinced pro-selfs, and 13 out of 36 pro-socials do not ignore when they could have. In both studies a substantial fraction remains for behavior that is neither always selfish, nor consistently other-regarding.

5.2. Alternative Explanations in The Literature

As shown in Section 4 subjects' choices in our experiment are consistent with cognitive dissonance. But what about other explanations brought forward in the related literature? Dana *et al.* (2006) suggest that beliefs about the receiver's beliefs drive their findings of dictator exit. As explained above such an effect of social norms is certainly plausible. The results in Tadelis (2007) indicate that shame may play a significant determinant of human behavior. Participants play a trust game variant and the experiment varies the extent with which the action of the trustee (cooperate or defect) is revealed. It is either i) kept private to the trustee, ii) revealed to the trustor, or iii) announced to the entire lab. Trustees cooperate significantly less often in the first condition than in the others, but there is no difference between the second and the third condition. Based on these findings shame would already have an effect, when only one other person is informed about an opportunistic action. Hence, shame could be an explanation for the difference between private setting and normal dictator game in Dana *et al.* (2006). In our design

---

[9]Besides the different setup (between- *vs.* within-subjects) slightly different payoffs were used in the two experiments.

the actual decisions of the dictators are not revealed (only outcomes). Shame should not matter and consequently no effect of negative self-related emotions (for instance, shame) is found in our data.

Dana *et al*. (2007) suggest that the higher implementation rates of a fair outcome in a transparent situation (baseline) compared to less transparent ones could be attributed to different behavioral norms these situations trigger. While acting generous may be the most relevant norm in the transparent dictator game, self-interest related norms gain relevance when some 'moral wiggle room' exists. Undoubtedly social norms are an important determinant for human behavior as illustrated by Bicchieri (2006). But individuals may be heterogeneous in how they *interpret the norm* for a given situation, as well as how they handle *deviating from this norm*. Our within-subjects design enables us to check whether—besides norms—individual-specific aspects play a role as well.

Our results do in fact indicate that *individual heterogeneity in how subjects respond to a deviation* from 'their' norm matters for behavior. When deviating, two psychologically conflicting cognitions ('their' norm and the deviant behavior) would create cognitive dissonance. Ignoring pro-socials behave according to their preferred behavioral norm under full information, but not anymore under 'strategic ignorance'. They experience cognitive dissonance, but the monetary gain outweighs that. In contrast, genuine pro-socials follow their norm in all situations, even under 'strategic ignorance'. If they deviated, the resulting psychological inconsistency would outweigh the monetary gain. Unstable pro-selfs experience cognitive dissonance in both situations since they do not follow their norm. But the monetary incentives outweigh this effect, leaving their choice unaffected. Convinced pro-selfs do not experience cognitive dissonance at a significant level, since the option that implies their norm and the one that yields the maximum monetary payoff coincide.

Of course, besides these aspects the *individual's interpretation of what the norm is* should also matter, and it should depend on the situation. For instance, ignoring pro-socials choose the fair option if the other's payoff is known but do not always uncover in periods 1/3. If the other's payoff is known their (expected) cognitive dissonance resulting from the selfish choice exceeds the monetary incentive. When the other's payoff can be ignored their (expected) cognitive dissonance does not exceed the monetary incentives of the selfish choice and in fact they do experience a significant conflict in periods 1/3.

5.3. *Discussion of Konow (2000)*

The analysis in Section 4 has shown that our data is largely in line with hypotheses 3.2 and 3.2 from Konow's model. A detailed analysis, however, shows that some discrepancies between the model and the data remain. In particular, to adhere to the *accountability principle*, specific assumptions on either α or β, or both, have to be made to explain the results. Suppose that the option to ignore the other's payoff reduces the costs of self-deception involved in reaching a certain ϕ. A lower β increases the optimal ϕ and y (see proposition 2C in Konow), *i.e.*, it results in a situation where the individual keeps a higher share but does not experience higher costs of self-deception or more cognitive dissonance. Applied to our experimental situation, a lower β could induce subjects categorized as *pro-social* in the open situation to ignore if given the opportunity. Interpreting the results, this means that 23 of the 36 *pro-social* subjects experienced a drop in β strong enough for them to change their decision. The remaining 13 experienced either no drop in β or it was too small to affect their choice. For the subjects categorized as *pro-self*, the situation is reversed. Since in the open situation they choose the corner

solution of the most selfish option, their choice cannot become less fair. However, a reduction in β leads to lower costs of self-deception given a certain choice. Hence, for all *pro-self* subjects who are sensitive to cognitive dissonance ($\alpha > 0$) and for whom the ignorance option leads to an at least marginal drop in β, ignorance is the dominant choice. For constant β, $\beta = 0$ or $\alpha = 0$, subjects are indifferent between ignorance and uncovering. Interpreting the results, this means that at least 23 of the 54 *pro-selfs* are either not sensitive to cognitive dissonance or self-deception, or experienced *no drop at all* in β. For the 31 who chose to ignore, cognitive dissonance weakly decreased since ϕ either increased (drop in β) or remained constant (constant β).

In addition to these somewhat specific conditions on α and β, the model predictions are not consistent with our cognitive dissonance proxies (the reported conflict and the decision time). For *unstable pro-selfs*, with ignorance the decision should become easier rather than harder, since the costs of self-deception (weakly) decrease. Hence, ignorance is either a dominant choice or subjects are indifferent (see hypothesis 3.2). In both cases, the decision should be simple. However, their decision times are significantly longer than those of genuine pro-socials (see Table 2).

Summarizing, the model of Konow (2000) predicts some main features of our data but has some difficulty explaining the details. This may be due to the somewhat restrictive assumption of the one-and-for-all behavioral standard implied by the accountability principle.[10] If subjects were allowed to hold different standards of behavior, the concept of cognitive dissonance may be better able to organize the behavior and proxy data we observe.

6. Conclusion

The extensive laboratory and field evidence on other-regarding behavior is usually explained with people having (genuine) social preferences. Recent experimental evidence on sorting and ignorance in allocation situations, including the evidence presented in this paper, has cast some doubt on this explanation. However, only few theory explaining this evidence has been put forward so far—with Konow (2000) as the notable exception.

We use the same theoretical framework as Konow's model—cognitive dissonance—to organize our data. The underlying assumption is that people do not show other-regarding behavior because they care for others, but because they want to maintain a certain self perception. Any deviating action from that self perception causes two psychologically inconsistent cognitions (their norm and the deviant behavior). Genuine pro-socials follow their behavioral norm in all situations and do not experience cognitive dissonance. In contrast, ignoring pro-socials behave according to their norm under full information, but not anymore when they can remain ignorant. They experience cognitive dissonance—a negative drive state—but the monetary gain outweighs that. The choice data of our within-subjects experiment is in line with this theoretical approach. We find a substantial extent of ignoring behavior among subjects. Further, our results indicate that such ignoring behavior corresponds to a higher experienced conflict and longer decision times (our proxies for the tradeoff between monetary utility and disutility from cognitive dissonance and self-deception), and less positive self-related emotions (our proxy for the negative drive state cognitive dissonance causes).

---

[10] Any common standard, be it derived from a single principle (as in Konow, 2000) or a combination of principles (as, e.g., in Konow, 2001 and 2010) may have similar problems explaining the data.

Thus, our paper adds to the experimental evidence on sorting and ignorance (Dana *et al.*, 2006; Lazear *et al.*, 2006; Broberg *et al.*, 2007; Dana *et al.*, 2007; Larson and Capra, 2009), connects this evidence to cognitive dissonance in the spirit of Konow (2000), and in addition successfully relates individual behavior to proxies of cognitive dissonance.

In spite of this supporting evidence, it must be emphasized that this paper is only one of a few explaining individual decision making with the concept of cognitive dissonance.[11] For the moment this rather small strand of literature around the seminal paper of Konow (2000) serves as an alternative to the known outcome-based and intentions-based social preferences models, and also to social approval models (see for instance Bénabou and Tirole (2006), Andreoni and Bernheim (2009) and Ellingsen and Johannesson (2009)). It remains for further research to systematically compare the cognitive dissonance-based approach to the more established social preferences models in order to find out more about the true determinants of human behavior.

Acknowledgements

References

1. Akerlof, G.A.; Dickens, W.T. The Economic Consequences of Cognitive Dissonance. *Amer. Econ. Rev.* **1982**, *72*, 307–19.

2. Andreoni, J.; Bernheim, B.D. Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica* **2009**, *77*, 1607–1636.

3. Aronson, E. The Return of the Repressed: Dissonance Theory Makes a Comeback. *Psychol. Inq.* **1992**, *3*, 301–311.

4. Beauvois, J.-L.; Joule, R.-V. *A Radical Theory of Dissonance. European Monographs in Social Psychology*; Taylor and Francis: New York, NY, USA, 1996.

5. Bénabou, R.; Tirole, J. Incentives and Pro-Social Behavior. *Amer. Econ. Rev.* **2006**, *96*, 1652–1678.

6. Bicchieri, C. *The Grammar of Society*; Cambridge University Press: New York, NY, USA, 2006.

7. Bolton, G.; Ockenfels, A. A Theory of Equity, Reciprocity and Competition. *Amer. Econ. Rev.*, **2000**, *100*, 166–193.

8. Branas-Garza, P.; Leon-Mejia, A.; Miller, L.M. Response Time under Monetary Incentives: The Ultimatum Game. *Jena Econ. Res. Pap.* **2007** (working paper).

9. Broberg, T.; Ellingsen, T.; Johannesson, M. Is Generosity Involuntary? *Econ. Lett.* **2007**, *94*, 32–37.

10. Camerer, C. *Behavioral Game Theory: Experiments in Strategic Interaction*; Princeton University Press: Princeton, NJ, USA, 2003.

---

[11]Note that also Houser *et al.* (2008) casually refer to cognitive dissonance to explain their experimental data.

11. Dana, J.; Cain, D.M.; Dawes, R.M. What You Don't Know Won't Hurt Me: Costly (But Quiet) Exit in Dictator Games. *Organ. Behav. Hum Decision Proc.* **2006**, *100*, 193–201.

12. Dana, J.; Weber, R.; Kuang, J.X. Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness. *Economic Theory* **2007**, *33*, 67–80.

13. Dufwenberg, M.; Kirchsteiger, G. A Theory of Sequential Reciprocity. *Game. Econ. Behav.* **2004**, *47*, 268–298.

14. Ellingsen, T.; Johanesson, M. Pride and Prejudice: The Human Side of Incentive Theory. *Amer. Econ. Rev.* **2009**, *98*, 990–1008.

15. Falk, A.; Fischbacher, U. A Theory of Reciprocity. *Game. Econ. Behav.* **2006**, *54*, 293–315.

16. Fehr, E.; Schmidt, K. A Theory of Fairness, Competition and Cooperation. *Quart. J. Econ.* **1999**, *114*, 817–868.

17. Fehr, E.; Schmidt, K. The economics of fairness, reciprocity and altruism—experimental evidence and new theories. In *Handbook of Reciprocity, Gift-Giving and Altruism*; Elsevier: Amsterdam, The Netherlands, 2006.

18. Festinger, L. *A Theory of Cognitive Dissonance*; Stanford University Press: Palo Alto, CA, USA, 1957.

19. Fischbacher, U. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Exp. Econ.* **2007**, *10*, 171–178.

20. Greiner, B. An online recruitment system for economic experiments. In *University of Cologne Working Paper Series*; University of Cologne: Cologne, Germany, 2004.

21. Harmon-Jones, E.; Mills, J. *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*; American Psychological Association: Washington, DC, USA, 1999.

22. Houser, D.; Xiao, E.; McCabe, K.; Smith, V. When Punishment Fails: Research on Sanctions, Intentions and Non-Cooperation. *Game. Econ. Behav.* **2008**, *62*, 509–532.

23. Konow, J. Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions. *Amer. Econ. Rev.* **2000**, *90*, 1072–1091.

24. Konow, J. Fair and Square: The Four Sides of Distributive Justice. *J. Econ. Behav. Organ.* **2001**, *46*, 137–164.

25. Konow, J. Mixed Feelings: Theories of and Evidence on Giving. *J. Public Econ.* **2010**, *94*, 279–297.

26. Larson, T.; Capra, C.M. Exploiting Moral Wiggle Room: Illusory Preference for Fairness? A Comment. *Judgment Decision Making* **2009**, *4*, 467–474.

27. Lazear, E.; Malmendier, U.; Weber, R. Sorting in Experiments with Application to Social Preferences. NBER working paper No. W12041; National Bureau of Economic Research: Cambridge, MA, USA, 2006.

28. Levine, D. Modeling Altruism and Spitefulness in Experiments. *Rev. Econ. Dyn.* **1998**, *1*, 593–622.

29. Oxoby, R. Attitudes and Allocations: Status, Cognitive Dissonance, and The Manipulation of Preferences. *J. Econ. Behav. Organ.* **2003**, *52*, 365–385.

30. Oxoby, R. Cognitive Dissonance, Status, and Growth of The Underclass. *Econ. J.* **2004**, *114*, 729–749.

31. Rabin, M. Incorporating Fairness into Game Theory and Economics. *Amer. Econ. Rev.* **1993**, *83*, 1281–1302.

32. Rabin, M. Cognitive Dissonance and Social Changes. *J. Econ. Behav. Organ.* **1994**, *23*, 177–194.

33. Rubinstein, A. Instinctive and Cognitive Resasoning: A Study of Response Times. *Econ. J.* **2007**, *117*, 1243–1259.

34. Steele, C.M. The psychology of self-affirmation: Sustaining the integrity of the self. In *Advances in Experimental Social Psychology*; Berkowitz, L., Ed.; Academic Press: New York, NY, USA, 1998; pp. 261–302.

35. Tadelis, S. The power of shame and the rationality of trust. In *University of California Working Paper Series*; University of California Berkeley: Berkeley, CA, USA, 2007 (mimeo).

Appendix

A. Experimental Instructions

Welcome and thanks for participating in this experiment.

In this experiment you can win a certain amount of money, which depends on your and the other participants' decisions in the experiment. It is, therefore, important that you read the following instructions carefully.

Please note that these instructions are only meant for you and that you are not allowed to exchange any information with the other participants. Similarly, during the entire experiment it is not allowed to talk to the other participants. If you have any questions or concerns, please raise your hand. We will answer your questions individually. Please do not ask your questions aloud. It is very important that you follow these rules, since otherwise we have to stop the experiment. Please also turn of your mobile phones now.

Procedure

The experiment lasts about 30 min. It consists of four parts. In each part you make decisions. Each decision will be explained again briefly on the monitor. While you make decisions, the other participants also make decisions which may influence your payoffs.

Your payoff from the experiment depends on your decision and the decisions of the other participants. But only one of the four parts is chosen randomly and you are paid in cash according to the payoff from this part. The exact procedure according to which your payoff is calculated is explained below. All amounts in the decision situation are given in EC (experimental currency). They are paid out in exactly the given amount at the end of the experiment. In addition you receive 2.50 Euro as a show-up fee.

After you filled in a questionnaire the experiment ends and you receive your payoff.

Again the procedure as an overview:

- Reading of the instructions, test questions

- Four parts with decision situations

- Questionnaire

- Payoffs and end of the experiment

Decision Situations

This experiment consists of four parts in which you make decisions. In each of these parts you are randomly matched with two other participants of the experiment. This means that your decision influences the payoffs of one other participant. In turn, the decisions of yet another participant influence your decisions. You can be sure not to meet the same participants twice.

Parts 1–3:

In the decision situations of parts 1-3 you can choose between five options. Each option includes a payoff for you and for one other participant. This participant is not known to you. In the experiment, he or she is simply called participant "Y". You can decide which option to choose, and with this choice you determine the payoff for yourself and participant Y.

In all decision situations, the payoffs 1, 1, 4, 4, and 7 are assigned to the options A, B, C, D and E. Only the numbers 1, 1, 4, 4, and 7 can occur as payoffs, both for yourself and for participant Y. Numbers/amounts assigned to options are determined randomly!

For example, assume that your and participant Y's payoffs are assigned to options according to the following Table (in EC):

Option	A	B	C	D	E
You	1	4	4	7	1
Participant Y	4	1	4	1	7

The choice of the options (A, B, C, D or E) that determines the payoffs for you and participant Y is up to you. Note that the above assignment of payoffs to options is only an example. It does not necessarily occur in the experiment.

It is important to understand that each participant receives his or her own randomly determined payoff table. With your choice of an option you determine the payoff for you and participant Y. In addition, another participant, who is called "X" in the experiment, determines the payoff for herself/himself and you through her/his own payoff table.

Which one of the two payoffs of each participant is paid out in the end—the one from his/her own decision or the one from the decision of his/her participant X, is determined randomly. The following diagram illustrates again the choice of the payoffs. Note that the participants X and Y you are matched to are different in each part of the experiment. This means you will not "meet" a given participant twice.

[A graphic that illustrated the matching of participants]

Hidden Payoffs

In some decision situations of parts 1-3 of the experiment you are informed about your own payoffs from the options A, B, C, D and E, but not about the payoffs of participant Y. On the screen you will see the following table:

Option	A	B	C	D	E
You	1	4	4	7	1
Participant Y	?	?	?	?	?

In this case you only know that the payoffs of participant Y are distributed uniformly. It is important that you understand what is meant by this.

As said above, in all decision situations the payoffs 1, 1, 4, 4, and 7 are assigned to the options A, B, C, D and E. Only the numbers 1, 1, 4, 4, and 7 can occur as payoffs. In each distribution 1, 1, 4, 4, and 7 occur for sure. This means twice the payoff 1, twice the payoff 4 and once the payoff 7. The combinations of options and payoffs are uniformly distributed, which means all occur with equal probability. As a consequence, for each option of your payoff table the payoff for participant Y amounts to 1 with probability 40%, to 4 with probability 40% and to 7 with probability 20%.

For example, two of the 30 possible payoff combinations for participant Y are:

Option	A	B	C	D	E
You	1	4	4	7	1
Participant Y	4	7	4	1	1

and

Option	A	B	C	D	E
You	1	4	4	7	1
Participant Y	1	4	7	1	4

Possibility to Uncover the Payoffs

If the payoffs of participant Y are hidden, e.g., you see the following table

Option	A	B	C	D	E
You	1	1	4	4	7
Participant Y	?	?	?	?	?

You have several possibilities:

- You can directly choose one of the options A, B, C, D or E.

- You can uncover the payoffs of participant Y. You can do this simply by clicking on the button "Participant Y" on the screen. You do not have to uncover, but can also choose an option directly. If you decide to uncover and click on "Participant Y", the payoffs of participant Y are shown. The assignment of your payoffs and the payoffs of participant Y does not change when you uncover. Neither does it cost anything. Afterwards you can choose one of the options A, B, C, D or E. By uncovering you neither receive new information about your own payoffs, nor about the payoff table of player X, the one who decides about your payoff.

- You can uncover the payoff table of participant X. You can do this simply by clicking on the button "Participant X" on the screen. You do not have to uncover, but can also choose an option directly. If you decide to uncover the payoff table of participant X, you see the payoffs of participant X and your own payoffs that participant X decides upon according to his payoff table. You cannot see which option participant X chooses! Uncovering the payoff table of participant X does not cost you anything either. You cannot receive any information regarding the payoffs of participant Y once you have chosen to uncover the payoff table of participant X. Afterwards you can just decide which of the options A, B, C, D or E to choose.

Part 4

In this part you also make decisions that are relevant for your own payoffs and the payoffs of participant Y. (Reminder: Participants X and Y are different in each part of the experiment.) For each decision you have to choose between two pairs of payoffs. Each pair of payoffs assigns one payoff to you and one to participant Y. You have to decide which of the payoff pairs you prefer. The particular pairs of payoffs you have to choose from are shown to you on the screen.

Overview of The Most Important Features of The Experiment:

- You make decisions in four parts of the experiment. In each part you are matched with two different participants X and Y.

- Your decisions also apply to the player who is participant Y in this part of the experiment.

- The player who is participant X in this part of the experiment makes his decisions (based on his own payoff table!), which assigns a payoff also to you.

- For the decision situations in parts 1-3 of the experiment you choose between five options. Your payoffs are uniformly distributed across options and always known to you. The payoffs of participant Y are uniformly distributed and not always known, but they can always be uncovered. You can choose one of the five options and with this choice determine the payoffs for yourself and participant Y.

- In part 4 of the experiment you choose between pairs of payoffs, that each assign one payoff to yourself and one to participant Y.

Your Cash Payoff from The Experiment

Your payoff from the experiment depends on your decision and the decisions of the other participants that you are matched to randomly in the four parts of the experiment. But only one of the four parts is chosen randomly and your are paid in cash according to the payoff from this part.

To determine this part and the decision that is relevant for your payment (yours or that of participant X), you have to make two more choices after part 4 of the experiment. For the first choice you choose between 8 letters. Each letter stands for one part (four parts) and one of the two "roles". With role we mean whether your decision or that of participant X is relevant for your payment. The order of the letters has no meaning related to the parts and roles. If you choose a letter which assigns to you a payoff from part four of the experiment, you have to make another decision (only then!). The reason is that in this part you make several decisions. For this second decision you again choose between letters, that is, you choose which of the decisions of part 4 is relevant for your payment. Again, the order of the letters has no meaning related to the decision that is relevant for you. In order for you to be able to check whether your choice is consistent with your payment, we provide an envelop with the assignment of letters to parts and roles and depose it well in sight. You can open it after the experiment.

With this procedure all parts of the experiment have the same chance of being relevant for your payment. In addition, the chance that your own decision is relevant for your payment is just as big as the chance that the decision of another participant is relevant for your payment. The same is true for all other participants.

Your payment is prepared while you complete the questionnaire. You receive your payment and sign a receipt before you leave the laboratory.

If you have questions regarding the content or procedure of the experiment, please raise your hand now. Then answer the following questions to ensure that you understood the instructions. Only if you answer these questions correctly you can participate in the experiment.

B. Post-Experimental Questionnaire

- First we would like to know how you felt in the situations during the experiment where you had to decide whether you would like to receive the information about participant Y's payoffs or the payment table of participant X (periods 1 and 3). You see a list of words describing different

feelings and emotions. Please indicate how intensively you experienced each of them (not at all; a little bit; to some extent; pretty much; very much):

- happy
- angry
- guilty
- proud
- irritated
- ashamed
- determined
- confused

- Did you find it easy to decide in these situations, whether you want to receive information about the payoff of participant Y or the payoff table of participant X (very easy; rather easy; rather difficult; very difficult)?

- Did you find it easy to choose one of the options (very easy; rather easy; rather difficult; very difficult)?

- What has been your reason to view / not to view the payoff of participant Y?

- What has been your reason to view / not to view the payoff table of participant X?

- If you haven't viewed the payoffs of participant Y, which payoff did he/she receive as a consequence of your choice?

- In general: What kind of the behavior (choice of options) would you find appropriate in case you know all the payoffs (period 2)?

- What kind of the behavior is appropriate in case you don't know the payoffs of participant Y (periods 1 and 3)?

- Do you have further comments about the experiment? Suggestions? Criticism?

- Did you find the instructions easy to understand?

- Do you buy organic products (never; seldom; often; always)?

- Do you buy Fair Trade products (never; seldom; often; always)?

- Do you think it is important what others think of you (not important at all; hardly important; some importance; quite important; very important)?

- Do you think it is important to have principles and to act according to them (not important at all; hardly important; some importance; quite important; very important)?

- Think of periods 1 and 3 again. What do you think is the percentage of participants who decide to uncover the payoff of participant Y?

- In case your payoff from the experiment has been decided by participant X, do you think he/she looked at your payoffs before making a decision?