

Keskustelualoitteita #58  
Joensuun yliopisto, Taloustieteet

Do norms and procedures speak louder than outcomes?  
An explorative analysis of an exclusion game

Timo Tammi

ISBN 978-952-219-123-6  
ISSN 1795-7885  
no 58

# **Do norms and procedures speak louder than outcomes? An explorative analysis of an exclusion game\***

Timo Tammi

Faculty of Law, Economics and Business Administration  
University of Joensuu

## **Abstract**

Economists have recently produced theories and evidence concerning the role of social preferences in various circumstances of economic behaviour. This has meant that in addition to selfish preferences for one's own material payoffs, economists have become increasingly interested also in investigating one's preferences for material and nonmaterial payoffs of the opposing side, or the parties involved, in general. A recent branch in this research is built on the contractarian aspect of behaviour: are we willing to follow the rules we have collectively agreed upon? A case in point is the work by Sacconi and Faillo (2005) which shows that a remarkable portion of the players in an exclusion game shifted from a selfish strategy towards an equal division of monetary payoffs after they jointly agreed upon the sharing rule (of an equal division) and had good reasons to expect that their opposing players also agreed upon the rule. This paper studies the Sacconi-Faillo model by taking into account the players' process-regarding preferences in selecting the sharing rule. It also illustrates the measurement of procedural fairness with an analysis of empirical data gathered in a small-scale piloting experiment.

**Keywords:** conformist preferences, process-regarding preferences, behavioural economics, experimental economics

**JEL classification:** C7, C9

---

\* I would like to thank Marco Faillo, Ismo Linnosmaa and Niko Suhonen for their helpful comments on earlier versions of this paper.

## 1. Introduction

Let me predict the outcome of the systematic and comprehensive testing of behaviour in situations where self-interest and ethical values with wide verbal allegiance are in conflict. Much of the time, most of the time in fact, the self-interest theory...will win (Stigler 1981, 176).

During the last two decades economic theorists and experimental economists have produced empirical evidence and theories which cast, at least when interpreted broadly, serious doubts on Stigler's prediction above. Indeed, in most of the reports on experiments, self-interest motivation is outrun by various types of 'other-regarding' motivation, such as rewarding and punishing, contributing to a public good, as well as reciprocity and altruism. Recently, a more felicitous counterexample to Stigler's assertion is stated in terms of the conformist preference theory by Grimalda, Sacconi and Faillo (Grimalda and Sacconi 2002; Faillo and Sacconi 2007). The theory sees individuals as seeking compliance with a moral principle conditional to the belief that also other individuals seek similar compliance. The theory has implications for many real-world problems predominated by *incomplete and/or informal contracts*. These problems are often characterised as *social dilemma* situations typical in economic activities under labels such as education, health care, natural and recreational resources, public television, regulation, organizational performance and other uses of shared resources where the individual and collective interests are at odds.<sup>1</sup>

This paper contributes to the research agenda commenced by Faillo, Grimalda and Sacconi.<sup>2</sup> The major aim here is to show, if only tentatively, that, in situations where conformity to a norm is a crucial source of individual motivation, people may also value the fairness of the processes in which the norm is jointly agreed. The view adopted in this paper emphasizes the so called *voice*-dimension of collective decision-making. This refers to the extent to which those who are affected by a decision have an opportunity to contribute to, or be involved with, the decision-making procedure (Dolan *et al.* 2007; Anand 2001). Presumably, this dimension has a bearing on the fields and topics mentioned above in the context of the conformist preferences theory.<sup>3</sup> Another purpose of the paper is to study, in the context of the exclusion game, other factors of behaviour than conformity and fairness – namely, other-regarding preferences (altruistic motivations), reliance on the rationality of

---

<sup>1</sup> Conformist preference theory is applied to non-profit organization analysis by Grimalda and Sacconi (2005) and to corporate social responsibility problem by Sacconi (2007).

<sup>2</sup> See Grimalda and Sacconi (2005), Faillo and Sacconi (2007) and Sacconi and Grimalda (2007).

<sup>3</sup> Benz (2005) mentions (i) consumption, (ii) work and employment, (iii) democratic participation, (iv) public good allocation, (v) taxation, (vi) redistribution and inequality, (vii) organizations, and (viii) law as areas of empirical relevance.

others, and the esteem of independence of others' behaviour. To be sure the main interest is in a preliminary analysis and in piloting the devices that were applied in measuring these factors.

## 2. Theories of social, procedural and conformist preferences

### 2.1. Overview

In response to the outcomes in experiments where the behaviour of subjects has quite often deviated from the 'self-regarding preferences model' of standard economic theory<sup>4</sup>, economists have formulated theories of what they call social preferences. The following simplified scheme illustrates the overall conceptual structure of utility functions in these new theories:

$$\begin{aligned} \text{Utility} &\equiv \text{material payoff} \\ &+ s_1 \times (\text{aversion to disadvantageous difference in payoffs}) \\ &+ s_2 \times (\text{aversion to advantageous difference in payoffs}) \\ &+ s_3 \times (\text{kindness} \mid \text{expectation of the other's behaviour}) \\ &+ s_4 \times (\text{conformity to a shared norm} \mid \text{expectation of the other's conformity}). \end{aligned}$$

The scheme suggests that a person experiences utility, or derives pleasure, from the components in her utility function. In the utility theory's manner of speaking the person assigns real numbers (utils) to (i) items in her own material payoffs or consumption bundle, to (ii) the distance of her payoffs from the fair level of payoffs, to (iii) kindness (or maliciousness), conditional to her expectations considering the intentions of the other player(s), and to (iv) her conformity to an ideal or a shared norm, conditional to her belief of the conformity of the other player(s) to the ideal.<sup>5</sup> In addition, the scheme suggests that the impact of the other components of the function except the material payoffs are mediated by the corresponding sensitivity coefficient, labelled as  $s_1, \dots, s_4$  and in the scheme. This indicates that the strength of the impact of one particular component can vary between individuals and across situations. It may also take the value of zero, which means that in a particular situation, or for a particular person, the impact of that component in the utility function is zero. Accordingly, when a sensitivity coefficient becomes high enough, a person may choose to take another course of action than if she were motivated solely, or dominantly, by her self-interested material outcomes.

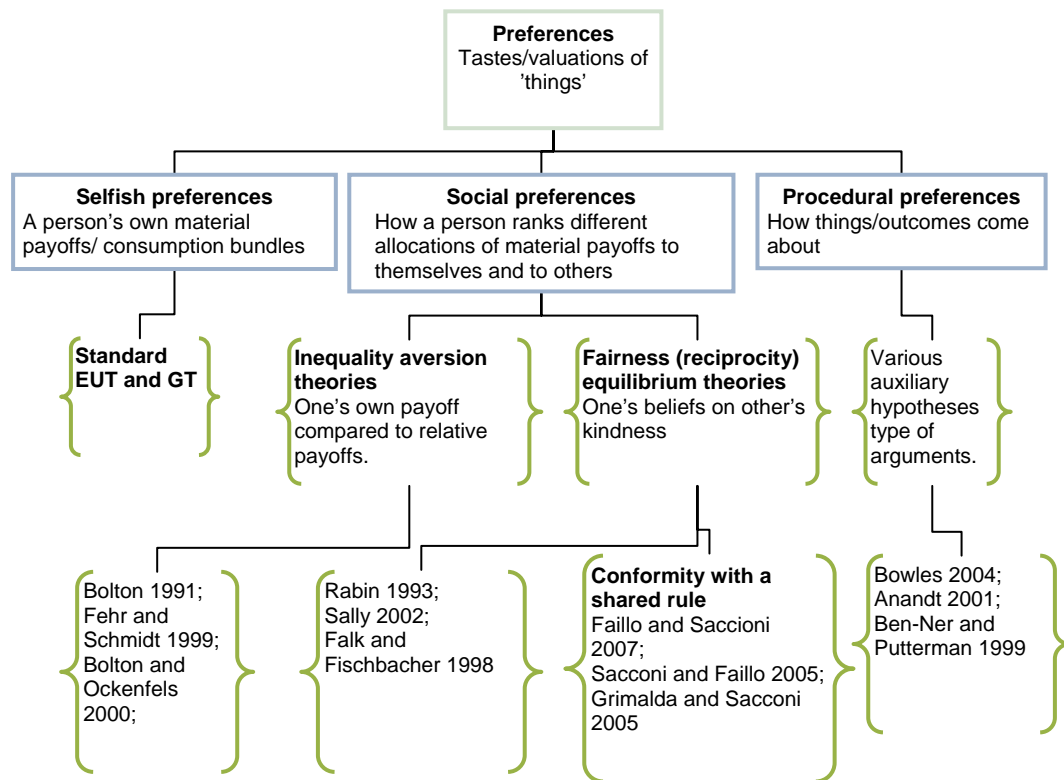
---

<sup>4</sup> The term 'self-regarding preferences model' comes from Cox (2004). The term makes it easier to see that experimental or other empirical findings which are inconsistent *with this model* need not be inconsistent with standard economic theory as utility functions can be defined as including others' payoffs in addition to one's own payoffs.

<sup>5</sup> Note that items (iii) and (iv) are modeled on the psychological games framework by Geanakoplos *et al.* (1989) where utilities are derived directly from beliefs.

In addition to the above scheme we can add a fifth component in the utility function by considering the fact that people often have process-regarding preferences. This means that people care how things and outcomes come about. A quite well-known demonstration of this is ‘Machina’s mom’ example where a mother who has one candy to be given either to her son or to daughter, and who decides to flip a coin to decide which one of the children will receive the candy (Machina 1989; 23-24; also Trautmann 2007). In the example the mother regards coin-flipping as a fair procedure of making a decision whereas the son sees it as unfair. As is the way with social preferences, the strength of the impact of process-regarding preferences varies between individuals and across situations.

Figure 1. Theories of preferences



**Some examples**

- tipping
- charity
- voting
- voluntary unpaid work
- working harder without money incentives than with money incentives
- contributing to public goods
- costly punishing of free-riders
- making generous offers in ultimatum games (and rejecting ungenerous offers)
- punishing and rewarding the opponent player according to his/her behaviour
- cooperation in a prisoner dilemma game
- punish those who violate a shared norm
- pure altruism (improving the wealth of others)

There is currently an array of theories of social preferences. Figure 1 is an attempt to organize these theories into a tentative typology according to the following criteria: (i) are preferences related to

one's own material payoffs, to the allocation of material payoffs, or to the process from which the payoffs come about; (ii) does a theory account for interactive expectations or not, and (3) do the preferences concern moral and ethical values or not. This typology helps us in understanding the main differences in the theories and in seeing the procedural and conformist preferences within a broader theoretical context.

Although a detailed account of the typology lies outside the scope of this paper, it can be noticed that all existing theories are based on methodological decisions to focus on particular components of utility and to ignore the other components. For example, reciprocity theories focus on material payoffs and reciprocity considerations while inequality aversion theories focus on material payoffs and differences in payoffs. A related fact is that so far as more psychology is brought in to these theories there emerges the dividing line between *fixed vs. constructed preferences*. That is, things such as fairness and norm compliance can be brought into the model of economic behaviour by assuming that people have 'ready-made' preferences towards these things. Thus, when people enter various decision-making situations they then consult exactly these ready-made preferences to make their decision. The standard expected utility theory and game theory, as well as many inequality aversion theories, operate just in this manner. On the contrary, Rabin's fairness equilibrium theory and the conformist preference theory operate by assuming that context matters. In other words, the theories assume that people have preferences that depend on what the others expect her to think or do (see, for instance Hargreaves Heap and Varoufakis 2004, ch 7 and Bruni and Sugden 2007).

An important fact is that experiments have had an important role in the development of the theorizing on social preferences. As has been argued by Sugden (2005) and Santos (2006), *these* experiments can be characterised as 'experiments as exhibits' aka 'behavioural experiments'. These characterisations point to the fact that typical experiments, where social preferences are investigated, include causal factors and mechanisms that *are not* 'built into' the theory that is currently tested (in addition to the factors and mechanisms which are included into the theory). Therefore, these experiments produce new knowledge on human behaviour and they result quite often in 'new findings', 'deviations', 'anomalies' etc. This capacity of behavioural experiments explains why there are currently several theories of social preferences as explanations of human behaviour which deviates from the predictions of the standard theory. The characteristics of behavioural experiments motivate also the study reported in this paper – viz., the operation of causal factors other than conformity in the context of the exclusion game.

The discussion on theoretical approaches and modelling principles continues to be vivid and also new theoretical ideas and extensions of theories and models appear almost all the time. Therefore

the above account excludes some of these new ideas. A recent hybrid model is Trautmann's (2007) *process fairness model* combining process-regarding preferences to the original Fehr-Schmidt model of inequality aversion. Another such model is presented by Charness and Rabin (2002) who combine social welfare preferences and reciprocity preferences with the player's material preferences to conceptually map the behavioural patterns found in experiments. A more radical view is that by Gold and Sugden (2007) which outlines a model for team reasoning in explaining some puzzles of game theory.

## 2.2. Process-regarding preferences

Like social preferences also process-regarding preferences have been known to economists for a long time, although they have been almost totally neglected in economists' theories and models. The term 'process-regarding preferences' is here taken to mean that *people do not only care about outcomes but they also, and sometimes mainly or even exclusively, value the processes and conditions where the outcomes come about*. For example, consider the Machina's mom example introduced above. In the example, the mother cares about the fairness of the process (of allocating a candy to one of her two kids) while the kid who was left without a candy considers the outcome fairness, and also, if complaining of the way the mother made her decision, also the process fairness<sup>6</sup>. The relevance of including process-regarding preferences into the theories of social preferences is apparent. Indeed, there is a 'shortage' of models that could explain the data that cannot be accommodated by inequality aversion theories or reciprocity theories (Sen 1997; 2000; Frey *et al.* 2003; Bolton *et al.* 2005; Trautmann 2007).

An illustrative example is provided by Lind *et al.* (1993; see also Frey *et al.* 2003), who investigated a situation where litigants faced a conciliation procedure in the end of which the court orders an award. The disputing parties can accept the award or reject it and go to trial. The litigants therefore balanced instrumental outcomes of the process against the fairness of the arbitration procedure. The result of the investigation was that the litigants who regarded the conciliation procedure as fair were more likely to accept the court-ordered award (*ibid.*).

---

<sup>6</sup> As for another example, consider Sen's (1997, 758-9) Chinese doctor case, where Dr. Chang, who works in a remote rural area in China and has one unit of medicine to be given either to sick child A or B. The medicine unit would save A's life with a slightly higher probability than it would save the life of child B. Dr. Chang's most preferred option is not to make a personal decision that would deny the medicine to one of the two children. What would be a proper mechanism or process through which a fair decision is to be made?

Another illustrative example of procedural fairness concerns the health care rationing. Think about priority setting decisions which are made by doctors and health care authorities and the outcomes of which influence patients and public. Doctors and authorities are here *strong actors* whose preferences and interaction determine the outcomes the patients, the *weak actors*, get from the operation of the health care organization. In Wailoo and Anand's (2005) survey research the respondents (who represented patients and public) agreed that procedures are important, especially they emphasized the idea that patients and public should be consulted before making health care priority settings. This suggests that the customers of health care organizations do not only care about the health consequences of the operation of the organizations but also about the fairness of the procedures from which the outcomes come about. By including into his view the idea of conformist preferences and ethical rules of taking care of also the weak actors in the society, we come to the following question: can the ideas of procedural fairness and conformist preferences be included into society's attempt to design institutions and organizations that reflect the preferences of all parties concerned? Furthermore, could the consideration of processes and conformity bring about better social benefits than institutions and organizations designed only by the strong actors?

Consider then a more abstract case called a three-person exclusion game *with* a sharing rule<sup>7</sup>. In a basic exclusion game two active players make their own individual decisions to take a share of the 'cake'. The passive third player gets the rest of the 'cake', possibly nothing. The situation raises the following considerations: would you take as much as possible without thinking the other two players; or would you guess what the other active player will do and do the same; or would you take less to yourself to ensure that the passive third player gets something? In the game *with a sharing rule*, the players *first* agree collectively, and without knowing their roles in the later phases in the game, on a rule according to which the problem of 'cake-sharing' should be solved. *Then* each player makes her individual decisions in situations where each player solves the 'cake-sharing' problem individually without any feedback on other's decisions, or possibility to signal or negotiate. As a player of such a game you probably would like to draw on the collectively agreed sharing rule, at least if you expect the others will do the same. This, indeed, is the *contractarian element* of the game: are you willing to follow the rule which you have jointly agreed with others?

A three-phase exclusion game combines the above sequences: *first*, playing the game without any explicit rule; *second*, collectively selecting a rule (and constituting a 'society'); and *third*, playing the game with the selected rule (in a 'society' just constituted). It is hypothesized that not only (i) a

---

<sup>7</sup> This example conceptually describes the experiment which is reported below. In the conceptual description we 'investigate' what causal factors would work in such a simplified situation while in the experiment we study, what will happen when real people behave in such a simplified but real situation.



player's beliefs of how the other players will behave (follow the collectively chosen rule or to deviate from it) influence on a player's behaviour but also (ii) her perception of the fairness of the process of collectively choosing the rule will affect on what she believes the others will do and on what she herself will do. As will be explained in section 2.3, the first influence is added to an individual's utility function in the conformist preferences theory by Grimalda and Sacconi. The second influence is preliminary explored in subsection 3.5.3.

The idea, that individual behaviour is affected by the perception of the fairness of the procedure by which a group makes its decision, needs to be theoretically justified. At least one should comment on the psychological plausibility of the argument. A short justification provided here refers to *the original instrumental value of a particular procedure which, however, has transformed during the (social) evolution of the procedure so as to have both an instrumental and inherent value, and even only an inherent value* (Anand 2001; see also Sen 1997, 749). For example, a coin-flipping method has its original value as a means to make decisions but it also has an inherent fairness-value.<sup>8</sup> The general point, however, is that the concept of procedural fairness is *observable and falsifiable*: its influences can be measured and it is possible to find situations where it does not play a role.<sup>9</sup>

### 2.3. Conformist preferences theory

In the context of social and economic interaction conformity refers to one's adjusting to norms, ideals, principles, or standards. That is, shared norms (and its counterparts mentioned) help us in situations where we do not solely, or at all, act to satisfy our desires. A telling example is the traffic rule saying that 'give way to traffic coming from the right'.<sup>10</sup> This rule, when shared by all drivers (and known by pedestrians) coordinates perfectly the traffic at crossroads. The point is that when you are in a *bounded rationality situation*, where you cannot use your computational power to form expectations of what the others will do (by regressing in the style 'think what others think you are thinking what they are thinking' and so on), the shared norm makes the forming of your expectation possible. Conformity with the norm is often a good response for all participants of the 'game' and it is often, but not always, secured by laws and informal rules with or without sanctions (see Bowles 2004, 47-48). Bicchieri (2006), for instance, argues that even if the conformity with a norm leads to choices that are against the person's self-interest, the person conforms to the norm if (i) she is aware

---

<sup>8</sup> For example, Pommerehne *et al.* (1997) studied people's attitudes towards various collective decision mechanisms in a hypothetical situation of placing a nuclear waste repository. They found that nearly 60% of the respondents were in favor to engineering mechanism and only 26% to a lottery mechanism.

<sup>9</sup> On a more profound account, see Benz (2005).

<sup>10</sup> This example comes from Hargreaves Heap (1992).

of the norm, (ii) she thinks a sufficient number of other players obey the rule, (iii) other players think the person ought to conform or a sufficient number of other players are ready to sanction her for not obeying the rule.<sup>11</sup>

In the traffic rule example above there are clear material benefits (the avoidance of a crash and its losses) which give a reason to a driver to conform the rule. The literature on why individuals deviate from the standard theory of rational choice, points, however, to the fact that very often the deviating behaviour is in contrast with the self-regarding material preferences of individuals. Accordingly, these theories typically focus on various *social dilemma situations*, where individuals' pursuit of self-interest is contradictory to the common interest of a group, organization or society. In this context the role of a rule is contractarian: which rule should be selected collectively, and how it should be selected in order to get people to cooperate for mutual benefit. If we assume that individuals are rational, each player of the game prefers the outcome where her ends are accepted by everyone; but since this would usually not lead to an agreement, individuals can be seen as willing to make concessions to reach agreement (Sugden 1993). Therefore, individuals would trade-off their own material payoffs against benefits from reaching agreement. These benefits may come from various sources, such as avoidance of penalties, enjoyment of solidarity, or pleasure of complying with a rule and so on.

The above account of some dimensions of norms gives a proper background to understand the so called conformist preference theory by Grimalda and Sacconi. Grimalda's and Sacconi's thrust for their theory is in combining a player's *intrinsic motivation* to act according to a shared principle with the player's *consequentialist* (selfish and material) preferences. They then construct what they call a comprehensive utility function where the utility depends on the player's material payoffs and the (dis)utility coming from the degree of one's conformity to a shared principle. Thus, the utility function consists of a *non-consequentialist aspect* in the sense that the preference to conform the shared ideal need not consult the player's preferences for her material payoffs. Instead, one's preference to conform is contextual – that is, it depends on what the other party expects her to think and do. In consequence, the conformist preference theory explains why subjects in an exclusion game experiment change their behaviour after they have jointly agreed a norm. This behavioural pattern is not explained by the inequality aversion theory or the fairness equilibrium theory.

Grimalda and Sacconi present the general form of their comprehensive utility function as follows<sup>12</sup>:

---

<sup>11</sup> Regarding the role of norms in the self-regulation by network-effects, see Aviram (2003).

<sup>12</sup> The presentation of the model follows Sacconi and Faillo (2005) and Faillo and Sacconi (2007).

$$V_i(\sigma) = U_i(\sigma) + \lambda_i F[T(\sigma)], \text{ where}$$

$\sigma$  = actions for the players;

$U_i(\sigma)$  = player's  $i$  material, consequentialist utility;

$\lambda_i F[T(\sigma)]$  = player's  $i$  ideal (non-consequentialist) utility which reflects  $i$ 's concern to the conformity with the shared norm in the function  $T$ ;

$\lambda_i \geq 0$  = weight parameter which expresses how important the conformist component is for  $i$ ;

$F$  = function which transforms social normative criterion into individual ideal utility;

$T$  = 'social welfare function' (takes a value for each state  $\sigma$ ).

The authors then present the definition of  $T$  with a Nash social welfare function. The idea in this definition is to provide 'metrics' for calculating the values of possible states of affairs. This results in an ordering of the possible states of affairs an impartial spectator would give on the basis of the social normative criterion (also Grimalda and Sacconi 2005). Furthermore, it is assumed that the ordering is shared by all actors who are involved into the 'game'. The exclusion game example makes this idea more concrete. There players select the social normative criterion (the sharing rule) in phase two. Let us simplify the selection by assuming that there are two rules, *high* ( $h$ , which means here 'as much as possible to two active players') and *low* ( $l$ , which means here 'equal shares to all three players the passive one included'). If the players jointly select the rule *low*, the rule *low* serves as the basis of the metrics calculation on the grounds of the Nash welfare function. This 'metrics' then produces the ordering  $N(l, l) > N(l, h) > N(h, h)$ , where the expressions ' $N(\cdot, \cdot)$ ' refer to the values the welfare function assigns to the active players' strategy choices. (This procedure, when applied to the exclusion game, yields the payoff matrix shown in Figure 4 in section 3.<sup>13</sup>)

The next step in constructing the model is the definition of the function  $F$  by using indexes of conformity to (two) players. **First**, *player  $i$ 's conditional conformity index* measures  $i$ 's degree of

---

<sup>13</sup> Formally,  $T$  is given by the Nash social welfare function  $N: T(\sigma) = N(U_1, \dots, U_N) = \prod_{i=1}^N (U_i - c_i)$ , where  $c_i$  is the disagreement utility resulting from the breakdown in the negotiation. In an exclusion game the two active players 1 and 2 ask either 'high',  $d_i^h = R/2$ , or 'low',  $d_i^l = R/3$ , where  $i = 1, 2$  and  $R$  is the amount of money to be divided in the game. Hence, the third player's payoff is  $s = R - (d_1 + d_2)$ , and, in particular, if both active players ask 'low',  $s = R - (d_1^l + d_2^l) = R/3$ . By assuming that  $c_i = 0$ , the Nash product yields the ordering  $d_1^l d_2^l s > d_1^h d_2^l s = d_1^l d_2^h s > d_1^h d_2^h s$ , based on

$$(i) T^{\max}(d_1^l, d_2^l) = N(d_1^l, d_2^l) = d_1^l \times d_2^l \times R/3,$$

$$(ii) T(d_1^h, d_2^l) = N(d_1^h, d_2^l) = d_1^h d_2^l \frac{s}{2} = N(d_1^l, d_2^h) = d_1^l \times d_2^h \times R/6, \text{ and}$$

$$(iii) T^{\min}(d_1^h, d_2^h) = N(d_1^h, d_2^h) = d_1^h \times d_2^h \times 0 = 0.$$

See the details in Sacconi and Faillo (2005).

deviation from the pure conditional conformity with the norm. The index, which varies from 0 (no deviation) to -1 (maximal deviation), is defined as follows (Sacconi and Faillo 2005):

$$f_i(\sigma_i, b_i^1) = \frac{T(\sigma_i, b_i^1) - T^{\max}(b_i^1)}{T^{\max}(b_i^1) - T^{\min}(b_i^1)}, \text{ where}$$

$b_i^1$  = player  $i$ 's belief concerning  $j$ 's choice;  $T^{\max}(b_i^1)$  = maximum attainable value of  $T$  given  $j$ 's choice according to  $i$ 's belief;  $T^{\min}(b_i^1)$  = minimum attainable value of  $T$  given  $j$ 's choice according to  $i$ 's belief; and  $T(\sigma_i, b_i^1)$  = actual value of  $T$  when  $i$  chooses strategy  $\sigma_i$  given her belief about  $j$ 's behaviour. In other words, the deviation index is composed of the *difference between* the value attached to  $i$ 's actual choice in the social welfare metrics *and* the highest value of the social welfare metrics calculated on the basis of collectively chosen rule *scaled* by the distance between the highest and the lowest welfare metrics values.

Although the selected social welfare criterion (the sharing rule) was assumed to be shared by all players of the game, the above formula of the conditionality of an actor's conformity adopts the perspective of an individual actor. This is just the idea that an actor's deviation from the social welfare criterion reflects how much she wants to generate the fulfillment of the outcome implied by the criterion, given the actor's beliefs of the other player's choice (see Grimalda and Sacconi 2005).

**Second**, player  $j$ 's expected reciprocal conformity index measures  $j$ 's degree of deviation from complete reciprocity in complying with the ideal principle  $T$ . This varies from 0 (no deviation at all) to -1 (maximal deviation). The index is seen through player  $i$ 's beliefs about  $j$ 's action and about her beliefs concerning  $i$ 's choice. The index is defined as follows:

$$\tilde{f}_j(b_i^1, b_i^2) = \frac{T(b_i^1, b_i^2) - T^{\max}(b_i^2)}{T^{\max}(b_i^2) - T^{\min}(b_i^2)}, \text{ where}$$

$b_i^1$  = player  $i$ 's first order belief about  $j$ 's action, which is formally identical to the strategy of player  $j$ ;  $b_i^2$  = player  $i$ 's second order belief about  $j$ 's belief about the action of  $i$ , which is formally identical to  $i$ 's strategy predicted by player  $j$ .

Also this formula presents an actor's perspective to the social welfare criterion. The formula shows how player  $i$  sees that the player  $j$  thinks about  $i$ 's conformity to the norm. This brings the principle

of mutuality to the model in that it assumes that an actor's motivation to comply with the norm increases the closer she thinks the other player complies with the norm.

The conditionality and reciprocity indexes above are then entered into the ideal component of the utility function. This is now defined in the following way:

$$\lambda_i[1 + \tilde{f}_j(b_i^2, b_i^1)][1 + f_i(\sigma_i, b_i^1)].$$

This definition states that if player  $i$  conforms perfectly to the ideal and if  $i$  expects that player  $j$  does the same, then the values of the two indexes are zero. In this case the resulting utility value of the ideal component is  $\lambda_i$ . In other words, the player's  $i$  utility value is the same as the importance of the ideal component to the player  $i$ . On the contrary, if the player  $i$  does not entirely conform and does not expect the player  $j$  to conform either, then both indexes take negative values, possibly -1, which yields a smaller utility value than  $\lambda_i$ .

Finally, the *comprehensive utility function*  $V_i$ , consisting of a material and an ideal component, can be written as follows (*ibid.*):

$$V_i(\sigma_i, b_i^1, b_i^2) = U_i(\sigma_i, b_i^1) + \lambda_i[1 + \tilde{f}_j(b_i^2, b_i^1)][1 + f_i(\sigma_i, b_i^1)].$$

This function implies that a player adopts the agreed rule if she expects that the other player acts in the same way. This also means that conformist preferences (utility from conforming) can induce players to choose strategies they would not choose if they consider their material utilities only.

The model can be summarised in the following way. In the beginning there is a *material game*<sup>14</sup> in which only the player's own material payoffs matter. This game is adequately captured by a utility function consisting only of a player's own consequentialist payoffs. Consider, then, the case that the game has a *social dilemma* structure: if the players would contribute to the common interest, then everyone would be better off than if everyone only pursues her own self-interest. Several new components can be added into the utility function to account for the altered structure of the game. The conformist preferences model alters the structure of the game in the following specific way. It assumes that an *ideal game* is involved with the material game in the sense that the interaction of strong players has influences on the weak players who, however, can not participate in the actual

---

<sup>14</sup> The terms 'material' and 'ideal' game come from Grimalda and Sacconi (2005).

game as active players. Therefore a *strong player's* description of the game now includes also a *weak player's* payoffs and the utility the *strong player* derives from the ideal component of the utility function defined above. In principle, the normative criterion  $T$  can be any principle of taking care of other players, especially the weak ones. The model, however, makes a 'stylised' case by assuming that the players of the game commonly, and as impartial spectators, select a rule to be shared by all players of the game. In addition, not only the compliance to the rule but also players' expectations of others' compliance enter the picture as sources of utility.

Recall now the idea of process-regarding preferences introduced in section 2.2. Clearly, such preferences reflect the players' acceptance of the procedure by which the selection of the rule was accomplished. In addition, they may have influences on the players' behaviour and beliefs and it can be hypothesized that if a player sees that the selection of the rule was fair (in that it reflects her preferences) it is more likely that the player conforms to the rule and believes that the other players also conform to the rule. This hypothesis is studied in section 3.5.3 below.

### 3. Analysis of an explorative experiment

#### 3.1. The outline of the Sacconi and Faillo (2005) experiment and the predictions of the model

Sacconi and Faillo (2005) conducted an exclusion game experiment to assess empirically the conformist preferences model. The experiment consisted of three phases in the following way: *in phase one* the subjects played three rounds of exclusion game with a reassignment of the player roles in the beginning of each round; *in phase two* the subjects were divided into three-member groups and each group voted the sharing rule according to which the sum of money (12 euros) should be divided; *in phase three* the exclusion game was again played in two rounds and the active players made public predictions of the other player's behaviour just before their own choice.

Figure 2. The exclusion game payoff matrix (Sacconi and Faillo 2005)

		G2		
		3	4	6
G1	3	3, 3, (6)	3, 4, (5)	3, 6, (3)
	4	4, 3, (5)	4, 4, (4)	4, 6, (2)
	6	6, 3, (3)	6, 4, (2)	6, 6, (0)

The experimental exclusion game payoff matrix is shown in Figure 2. Two active players (G1 and G2) make their own decisions and the passive player (G3) gets the sum which depends on what the active players decide.

Figure 3. Sharing principles and rules in the SF experiment (*ibid.*)

Principles	Principle 1 Every player should share the benefits, in particular, who has not the possibility to choose should not receive less than the others.			Principle 2. People who play under a decisional role could claim a higher share of benefits.				
Rules		G1	G2	G3		G1	G2	G3
	1.1	33 %	25 %	42 %	2.1	50 %	33 %	17 %
	1.2	25 %	33 %	42 %	2.2	33 %	50 %	17 %
	1.3	33 %	33 %	33 %	2.3	50 %	50 %	0 %

In voting the sharing rule the subjects voted one of two general principles and among six specific sharing rules derived from the principles (Figure 3.). In the voting procedure the anonymity was maintained and no communication among subjects was allowed.

Figure 4. Payoff matrix when the conformist preference model is applied (*ibid.*)

		G2		
		3	4	6
G1	3	3, 3	$3 + \frac{3}{4}\lambda_1, 4 + \frac{3}{2}\lambda_2$	3, 6
	4	$4 + \frac{3}{2}\lambda_1, 3 + \frac{3}{4}\lambda_2$	$4 + \lambda_1, 4 + \lambda_2$	4, 6
	6	6, 3	6, 4	6, 6

Figure 4 is based on the conformist preferences model and the social welfare metrics explained in section 2.3. It shows the individual comprehensive utility values for the exclusion game payoff matrix in Figure 2, on the condition that in each state the player's beliefs reciprocally predict the strategy chosen by the opponent. The diagnosis of the game goes now as follows:

- (1) if  $\lambda_1 > 6 - 4 = 2$  then player one prefers, in terms of her comprehensive utility function, strategy 'ask for 4' to strategy 'ask for 6' as well as player two; we can define that the combination 'ask for 4', 'ask for 4' is a psychological equilibrium if  $\lambda_1 > 2$ ;

- (2) if player one believes that player two does not choose according to the shared principle, she does the same and player two decides symmetrically so that both choose ‘ask for 6’; we can define that the combination ‘ask for 6’, ‘ask for 6’ is also a psychological equilibrium.

Sacconi and Faillo then produce the following predictions of their model<sup>15</sup>:

- (1) Because the possibly ‘innate’ conformist preferences are inactive in phase one of the game almost all players choose ‘ask for 6’;
- (2) A significant part of subjects choose the sharing rule that assigns equal payoffs to all three players;
- (3) After having selected the sharing rule, a significant number of players generate expectation of reciprocal conformity and choose according to it;
- (4) A significant part of players choose ‘ask for 4’;
- (5) Information that confirms beliefs about conformity of others does not change the willingness to conform.

Sacconi and Faillo found evidence in support of the conformist preferences model. As they argue, not only the fact that the players changed their behaviour between the phases one and three, but also the fact that many players expressed the same strategy choice between the phases, is consistent with their theory (given their choice in phase two and their beliefs). In other words, the latter fact does not serve as an anomaly of the theory. But, as we shall see, this is not the whole story. The introduction of different collective decision-making mechanisms induces some differences in the behaviour of subjects in the experiment.

### **3.2. An explorative experiment: design and its evaluation**

The explorative experiment, which this paper is based on, was conducted in January 2008 at the University of Joensuu (hereafter J08 experiment). The design of the experiment reproduces the main structure of the Sacconi-Faillo experiment (hereafter SF experiment) but it also extends it by varying the procedures with the help of which the collective decisions were carried out.

---

<sup>15</sup> As noticed in section 2.3, the inequality aversion theory by Fehr and Schmidt and the fairness equilibrium theory by Rabin do not predict correctly in the context of the exclusion game. That is, both of them predict (although for different reasons), that the players choose either ‘ask 4’ or ‘ask 6’ both in phase one and in phase three.



Figure 5. Outline of SF and J08 experiments and of their comparison

SF experiment:  
J08 experiment

NrOXrO  
O<sub>b</sub> NrOXrO NrOXrO NrOXrO O<sub>a</sub>

In Figure 5 above N refers to non-randomized assignment of subjects to the treatment group (since there is only the treatment group). The Os in the figure indicate that certain variables are measured both before the treatment (X) and after the treatment. The small letters ‘r’, in turn, refer to the fact that the constitution of the sub-groups was randomised before each observation. The symbols O<sub>b</sub> and O<sub>a</sub> refer to measurements done before and after the actual exclusion game experiment. In SF experiment there was one treatment and one measurement before and one measurement after the treatment. The J08 experiment, on the other hand, adds two more treatments. In addition, it incorporates measurements before the exclusion game experiments and after these. The whole experiment consisted of three stages and stage two consisted of three sessions (Figure 6.)

Figure 6. Structure of the experiment

Stage one	Stage two			Stage three
	Session one Lottery	Session two Majority voting	Session three Negotiation	
O	N O X O	N O X O	N O X O	O
Dictator game Secure game	Three rounds before the treatment	As in session one	As in session one and two	Evaluation tasks
	Three rounds after the treatment			

It is also worth of noticing that none of the three treatments in J08 are identical to the SF treatment. In other words, when SF employed the three-person group unanimity voting procedure in selecting the sharing rule, J08 made use of the following three procedures: (i) lottery, (ii) a whole-group majority voting and (iii) a three-person group negotiation<sup>16</sup>.

The design allows the construction of several derived variables, such as distribution of payoffs, differences between choices before and after the treatment, and differences between the sharing rule payoff and a players’ choice. Although the sample size is small (18 subjects/36 – 54 choices) due to the preliminary nature of the experiment, some qualified results can be reported from the

<sup>16</sup> Before negotiation, the subjects were divided into three-person groups. Each group negotiated within the group and selected the sharing rule. No information on the outcomes of the negotiation was distributed to other groups. After the negotiation, the same groups played the exclusion game and the roles of the players within the group were changed between each round.

experiment. Although exact statistical tests ‘correct’ the biases the sparse and unbalanced data may cause in the use of asymptotic tests, it can not be ruled out that another sample of subjects would possibly produce different results.

The experimental design raises the question of unwanted *maturation effects*. This means that when the subjects enter in session two (and three, respectively) they already are affected by the treatment in session one (and two). Table 1 shows that the possibility of these effects cannot be ruled out: the choices labelled *E* (‘equity-oriented’, that is, choice values up to 4) progress in an ascending way in before treatments choices 1, 2, and 3, while choices labeled *S* (‘self-interest-oriented’ choices, namely choices 5 and 6) progress in a descending way. However, the Cochran exact test sees no statistically significant differences in the before-treatment choices between the sessions ( $p = .148$ ). The same holds for the after treatment choices ( $p = .841$ ). On the other hand, the chi-square test shows that although the choice distributions in the after-treatment choices in session one and in the before-treatment choices in session two are different and skew to opposite directions, the difference is not statistically significant ( $p = .061$ ). The same holds between the after-treatment choices in session two and the before-treatment choices in session three ( $p = .236$ ). Since maturing effects are obvious, at least in the sense that the end of a session is ‘too similar to’ the beginning of the next session, we must keep in mind that these effects may increase the occurrences of the behavioral patterns predicted by the conformist preferences theory. At the same time, the maturing effects may obscure the effect-size that can be leveled against each treatment.

Table 1. Evaluation of the maturing effects

	Frequencies		<i>p</i> -values			
	E	S	Chi-Square Exact .061	Chi-Square Exact .236	Cochran Exact .148	Cochran Exact .841
Before treatment 1	12	24			X	
After treatment 1	21	15	X			X
Before treatment 2	15	21	X		X	
After treatment 2	21	15		X		X
Before treatment 3	17	19		X	X	
After treatment 3	19	17				X

It is also worth of noticing that the treatment in session three (group-negotiation) was substantially different from the other two treatments. The treatment allowed that each group selected a sharing rule of its own. In consequence, there emerged selections of the ‘active players divide the whole cake’ rule, which meant that also the number of choices labelled ‘*S*’ increased from the conformist reasons and not (maybe) from the maturing effects.

### 3.3. Hypotheses

The main purpose of the experiment was to find out whether the different collective decision-making procedures have different influences on the behaviour of subjects. As was suggested in the theoretical section of the paper, it is reasonable to argue that different procedures work differently and give birth to different situation-specific perceptions of fairness. Hence, we can set the following two hypotheses:

- (1) Although the subjects still exhibit conformist motivations, there are differences in how many decisions are changed when moved from the ‘before treatment’ to the ‘after treatment’ phase;
- (2) The subjects’ perceptions of the fairness of the selection procedures matter; in particular, if a subject sees that her own preferences are taken into account, she is more likely to conform to the selected rule than otherwise.

### 3.4. Experimental procedures

The subject pool of the experiment consisted of the students in the faculty of Business, Law and Economics at the University of Joensuu. The subjects, 18 in total, were recruited by an e-mail message. Subjects were paid 5 euros for participation and up to ten euros according to their decisions in the experiment. The experiment sessions were run manually and the whole experiment lasted approximately 45 minutes. Subjects were given the instructions (including the information in Figure 2 and Figure 3) before the experiment. The progress of the experiment was conducted by the experimenter by using slides and by distributing separate answering sheets in stages.

The first stage of the experiment consisted of a *dictator game* and a *secure game* (see Appendix 1.). In each main session in the experiment a subject played the exclusion game for six rounds – three rounds before the treatment and three rounds after the treatment. The total number of rounds in stage two was eighteen. The subjects were told, that for each subject one round was selected at random and subject earnings were derived from it. Regarding stage one, it was told that for each task separately, one of the subjects was selected at random for paying for real.

In the exclusion game sessions the subjects were assigned to three-member groups before every three-round set. The roles of the players changed so that each player took each of the roles G1, G2, and G3 in turn. In five out of six three-round sets the group membership was anonymous. That is,

although the subject knew her group label in each three-round set, she didn't know who the other members in the group were. In the last three-round set the subjects knew the other members in the group as the treatment was based on face-to-face negotiation within each group.

The treatments in stage two were conducted in the following way. In *the first session* the subjects were told that one of the six sharing rules in Figure 3 will be selected at random by throwing a dice. The dice was thrown by the experimenter, and the outcome was announced for the subjects. In *the second session* the subjects were invited to vote for the sharing rule. Before voting it was told that the 50% majority rule will be applied and that new rounds of voting will be conducted until the result fulfils the majority condition. The voting was carried out by a secret ballot procedure and the winning rule, as well as the scores of the alternatives, was announced to the subjects. In *the last treatment* based of group-negotiation the subjects were first given their group symbols and then invited to stand up and to search the group co-members and to negotiate and select the sharing rule together. Table 2 shows the distribution of votes in the second session and the distribution of the group-negotiated selections in the third session.

Table 2. Votes in session two and selected rules in session three

Sharing rule (figure 3)	Session two: majority voting		Session three: group negotiation	
	Frequency	Percent	Frequency	Percent
1 (1.1.)	-	-	-	-
2 (1.2.)	1	5,6	-	-
3 (1.3)	10	55,6	15	83,3
4 (2.1.)	-	-	-	-
5 (2.2.)	1	5,6	-	-
6 (2.3.)	6	33,3	3	16,7
Total	18	100	18	100

### 3.5. Results

#### 3.5.1 Choice distributions

The J08 experiment replicated the major outcomes of the SF experiment. The bar charts in Figure 7 show that the distributions of payoffs changed both when moved within the session (from decisions before to decisions after the treatment) and between the sessions. We can see that in all rounds before the treatment the *mode value* was six (euros) but after the treatment four (euros). This indicates, if only at a very general level, that the shift in the behaviour of players within each session was to the direction predicted by the conformist preference theory. It is worth of noticing, however,

that the bar charts below include also the payoffs of the passive players. These are accumulated on the low values of horizontal axis and also on value four but not on value six.

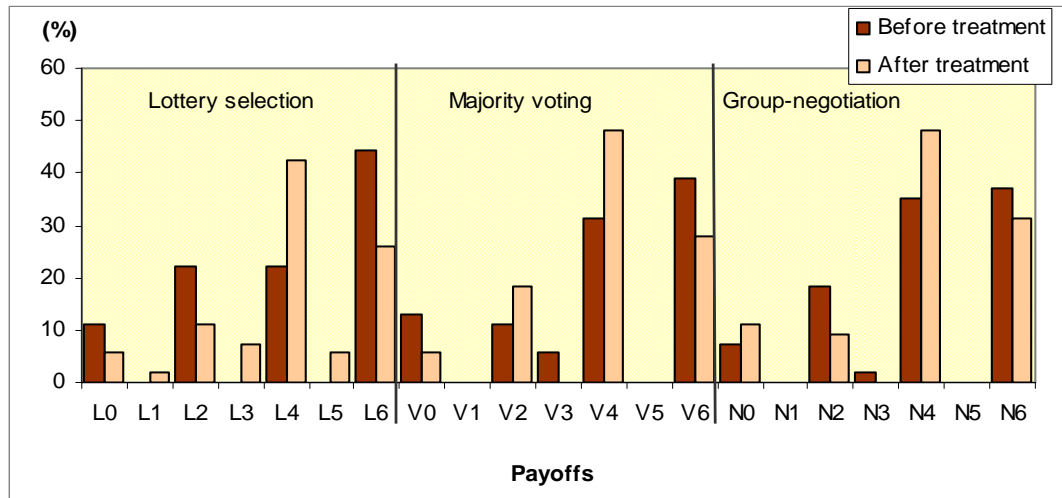


Figure 7. Players' choices in the experiment

### 3.5.2 Behavioural patterns and effect-size evaluation

Figure 8 shows the percentage values of the after-treatment behavioural patterns in different sessions. Our focus is here on conformist patterns, which are defined as rule-choice-expectation patterns (*RCEs*, in short) that fulfil the following condition:

$$R = C = E^{17},$$

where  $R$  refers to the payoff implied by the agreed rule,  $C$  to the payoff related to the strategy chosen and  $E$  to the other player's payoff due to her expected action. Defined in this way, there are two conformist patterns in our data. The *first one* occurs when the jointly agreed rule is the 'equal-split' rule and a player's strategy choice and expectation coincide with the rule. The *second conformist* pattern occurs, when the players agree on the rule called 'two active players divide the whole cake', and a player's strategy choice and expectation coincide with the rule. Let us label the former one  $R_{es}$  and the latter one  $R_{us}$ . In the J08 experiment, the  $R_{us} = C = E$  patterns occurred only in session 3 (with the group-negotiation procedure).

<sup>17</sup> One can read this 'conformer's reasoning' in the following way: "My choice pays me out the same amount of money that pays out the choice I expect my opponent player will make and, in addition, this amount of money equals the payoff implied by the sharing rule I and my co-player(s) have selected".

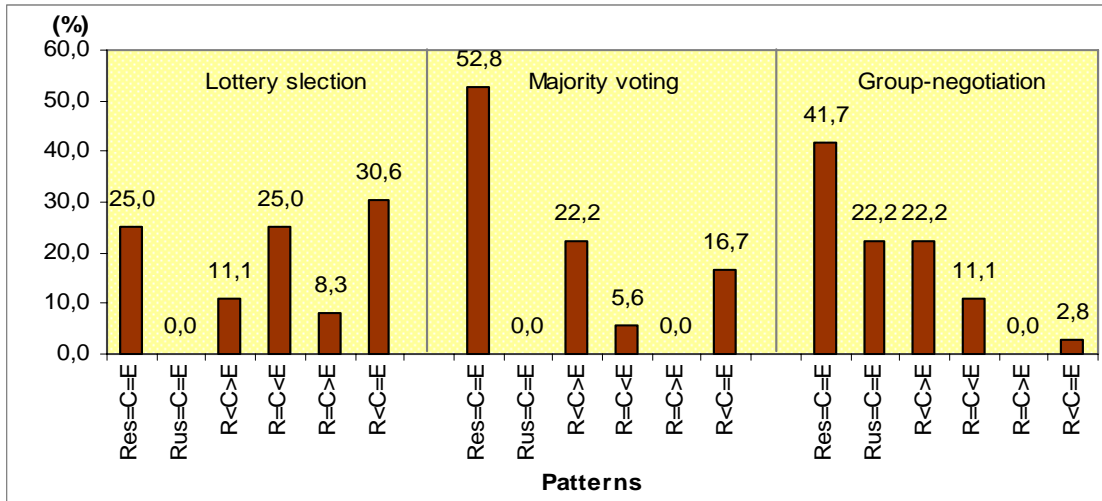


Figure 8. RCE-patterns in different sessions

One main concern of the experimental analysis was the number of decisions that changed between the ‘before-treatment’ and ‘after-treatment’ choices. Related to this, the first hypothesis in section 3.3 suggested that there are differences between the treatments. The McNemar test was performed to study the difference between before-treatment and after-treatment behaviour. The difference is significant only in the case of the lottery-treatment ( $p = .006$ ). The disappearance of the significance in the other two treatments is a likely result of the unwanted maturation effects mentioned in section 3.2. Accordingly, as Figure 9 shows, the number of equity-oriented choices increases in the before-treatment phases when moving from session one to sessions two and three.

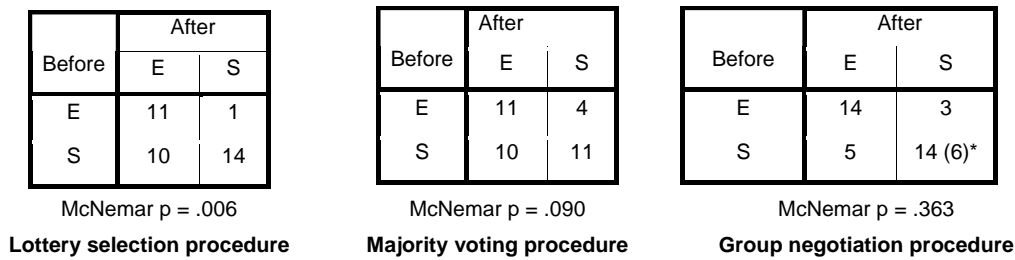
A straightforward (but statistically debatable) interpretation of the McNemar test indicates the following *effect-sizes*<sup>18</sup>:

Lottery treatment:	41,7 (%)
Majority voting treatment:	47,6 (%)
Group negotiation:	62,5 (%)

The effect sizes above should be taken with reservations, since the obvious maturing effects decrease the row sum of cells *c* and *d* (see footnote 18). Note also, that the group-negotiation procedure allowed the agreement on the group-specific sharing rule. Indeed, one group agreed on the rule 2.3 (the active players divide the cake and the passive player gets nothing). The choices of this group are removed from the effect-size calculation.

<sup>18</sup> There is no standard measurement of the effect-size in McNemar test. The effect-size  $e$  is here calculated from the formula  $e = c/(c + d) \times 100$  when the 2x2 table is:

<i>a</i>	<i>b</i>
<i>c</i>	<i>d</i>



\*The figure in the brackets shows the number of S-choices by players who agreed on the 2.3. rule in this session.

Figure 9. Comparison of equity-oriented (*E*) and self-interest-oriented (*S*) distributions of choices before and after the sharing rule selection

Taken together, the analysis of conformist patterns and the evaluation of the effect-sizes suggest that the portion of conformist patterns increased when moving from session one to sessions two and three. In addition, the effect-size increased respectively (although only symptomatically). There can be (at least) two reasons to this: maturation effects and procedural fairness considerations. Maturation effects are obvious, especially the effects of information spreading, since the majority voting procedure revealed information on how the votes on the sharing rules were distributed. As will be shown in subsection 3.5.4, the relation between the behaviour in the dictator game (and the secure game) and the behaviour in the exclusion game changed substantially after the completion of the majority voting procedure. The experimental design does not allow, however, a closer analysis of the information effects. Instead, the influence of fairness perceptions is analyzed below.

### 3.5.3 The influence of fairness perceptions

The second hypothesis in section 3.3 suggests that if a player sees that the procedure of jointly selecting the sharing rule is fair, she is more likely to exhibit conformist preferences than otherwise. This means that a player's fairness perception influences on what she believes the others will do and on what she herself will do. This was investigated by measuring the subjects' perceptions by two devices – viz. by a questionnaire type Likert scale measurement and by the so called *IOS scale* (Inclusion of Other in the Self scale). The Likert-scale measurement was applied by asking the subjects to indicate whether they think a given procedure was 'very fair', 'fair', 'unfair', or 'very unfair'. The items which they were asked to rank were different decision procedures in general (lottery, negotiation, majority voting, and expert system) and the decision procedures used in the experiment. The IOS-scale, on the other hand, measures the subject's conception of her standing relative to the other participants in the decision-making (see De Cremer *et al.* 2005). The method is based on the idea that a subject is given different descriptions of her standing with the help of

variously adjusted circles, where one circle represents the subject and the other circles stand for the others (see Appendix 2). The subject then selects the description which corresponds with her conception. Regarding the experiment reported here, we can interpret that if the subject saw that she was an outsider in a group, she conceived the decision-making as unfair in the sense that her voice was not taken into account ('no-voice' category in the analysis reported below). On the other hand, if the subject saw that she was a member of the group, she thought the decision-making was fair as her voice was taken into account ('voice' category). This interpretation highlights the voice-dimension of procedural fairness in that it indicates the extent to which the actors who were affected by the collective decision saw they had an opportunity to contribute to the decision-making.

The majority of the subjects took a view that the various collective decision-making procedures are fair or very fair in general (Table A3.1 in Appendix 3). The same actually holds for the attitudes towards the procedures employed in the experiment except that here seven out of eighteen subjects regarded lottery mechanism as unfair or very unfair. An interesting notion was made in the statistical analysis (one-tailed Fisher's exact test) of the relationship between fairness rating and subject behaviour. Namely, those who rated voting mechanism and group-negotiation as *generally* fair or very fair, exhibited very often equity-oriented choices in corresponding after-treatment phase in the experiment; respectively, those who rated these mechanisms as generally unfair or very unfair, exhibited *only* self-interest-oriented choices ( $p = .003$  in the voting treatment and  $p = .006$  in the group-negotiation treatment). No other significant relationships, however, were found regarding the Likert-scale rankings. This needs, of course, closer scrutiny in the future experiments.

Table 3. Fairness perception and the deviation of choices from the agreed rule

	Majority voting				Group negotiation			
	Conforms	Deviates	Total	$p^*$	Conforms	Deviates	Total	$p^*$
No voice	4	12	16	.001	6	8	14	.009
Voice	16	3	19		19	3	22	
Total	20	15	35		25	11	36	

\*Fisher's exact test

The IOS scale measurements produced preliminary and useful information on how procedural fairness perceptions may influence on the subjects' behaviour in the experiment (regarding frequencies, see Table A3.2, Appendix 3). Accordingly, Table 3 above shows how many 'no voice' vs. 'voice' choices conformed to or deviated from the collectively selected sharing rule. As can be seen, the 'voice' choices conform more often to the rule than 'no voice' choices ( $p = .001$  in the majority voting treatment and  $p = .009$  in the group-negotiation treatment; one-tailed Fisher's exact test).



The analysis shows also that one's perception of the fairness of the decision-making was related to what one believed the other parties in the game will do. A one-tailed Fisher's exact test was performed to study the relation between one's perception of the situation and one's prediction of what the other players will do. As shown in Table 4, the relation between these variables was significant in the majority voting treatment ( $p = .018$ ) and in the group-negotiation treatment ( $p = .013$ ). Those who conceived their voice was taken into account predicted more often than the others that her opponent players conform to the jointly agreed rule.

Table 4. Fairness perception and the deviation of one's beliefs (of others' choice) from the agreed rule

	Majority voting				Group negotiation			
	Conforms	Deviates	Total	$p^*$	Conforms	Deviates	Total	$p^*$
No voice	15	9	24	.018	13	8	21	.013
Voice	27	3	30		30	3	33	
Total	42	12	54		43	11	54	

\*Fisher's exact test

Finally, it is interesting to see whether there is a relationship between fairness perception and *RCE*-patterns. This was studied by performing (again) a one-tailed Fisher's exact test. The relation was significant in the majority voting treatment ( $p = .014$ ) and in the group-negotiation treatment ( $p = .005$ ). Table 5 shows that if the subject thought her voice was taken into account in the collective decision, she produced more often conformist *RCA*-patterns than otherwise. Respectively, if the subject saw that her voice was not taken into account, she produced non-conformist patterns more often than otherwise. This suggests an interpretation that the 'voice'-condition is an important and influential dimension in procedural fairness: if the players see that they were participants of the social choice, they are very likely to conform to the selected rule and expect that also others conform to the rule.

Table 5. Fairness perception and the *RCE*-patterns after the majority voting procedure and after the group-negotiation procedure

	Majority voting				Group negotiation			
	Conformist	Non-conformist	Total	$p^*$	Conformist	Non-conformist	Total	$p^*$
No voice	5	11	16	.014	4	10	14	.005
Voice	14	5	19		17	5	22	
Total	19	16	35		21	15	36	

\*Fisher's exact test

We can conclude, if only on the basis of a small-scale data, that in the experiment *players' perception of the fairness of the collective decision making procedure had an influence on their*

behaviour and on their beliefs about the other players' behaviour. In particular, if a player thinks that her preferences are reflected in the collective choice of the sharing rule, it is likely that she behaves in the way predicted by the conformist preference theory. The result is similar to the one found by DeCramer et al. (2005, 401) in a different, but related context of cooperation managing. These results give support to the idea presented in section 2.2 that economic theories of human behaviour should pay more attention to the questions and aspects of procedural fairness. Regarding the conformist preference theory, the outcome encourages the investigation of an additional component of fairness to be integrated into the theory.

### 3.5.4 Selected observations of the dictator game and the secure game

The experiment reported here consisted also of two tasks that are of secondary interest in this paper. First of the tasks was a 'choosing to take' version of a dictator game (Rubinstein 1999) and the other was the game labelled as 'relying on other's rationality' also called a 'secure' game (*ibid.*). In a dictator game a player is asked to tell the sum of money she chooses to take herself in dividing ten euros between herself and another player. In a two-player secure game player 1 chooses between two actions A and B: action A yields the outcome (5; 5) if player 2 chooses A and (0; 4) if player 2 chooses B; respectively, the outcomes are (2; 2) and (3; 3) if player 1 chooses B (see Appendix 1). We here study tentatively, whether there is any relationship between the behaviour of subjects in these games and in the exclusion game.

Figure 10. Choices in the dictator and secure games

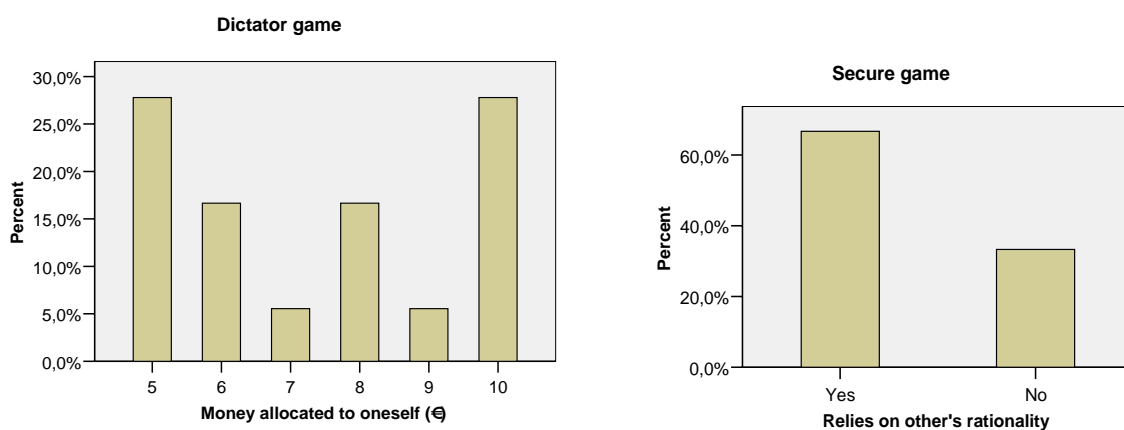


Figure 10 shows the choices in the dictator game and secure game. In the case of the dictator game the symmetry of the percentage distribution of choices is outstanding: proportions of equal split (5; 5) and absolutely unequal split (10; 0) are even, and the middle-value 7,5 divides the observations in half. We can interpret that those players who chose to take a *medium share* (from 5 to 7 euros) of the

whole sum had a stronger other-regarding motivation than those who chose to take a *large share* (from 8 to 10 euros). A Fisher's exact test was performed to study the relation between choices in the dictator game and the choices in the exclusion game. The relation was significant in choices after the lottery treatment and in choices before the voting treatment ( $p = .020$ , one-tailed Fisher's exact test in both cases). In these choices the medium share takers of the dictator game exhibited more often equity-oriented preferences in the exclusion game than did the large share takers. Respectively, the large share takers exhibited more often self-interest-oriented preferences than the medium share takers. Since the relationship disappeared in the later choices in the experiment, we can not reach any unambiguous interpretation. However, it can be suggested with reservations that (i) the subjects learned to play the exclusion game after the lottery treatment (as they found their other-regarding preferences) and that (ii) the nature of the game changed after information on other players' preferences was distributed in the majority voting treatment. Again, additional experiments with proper designs are called for.

Regarding the secure game, the proportion of those who exhibited non-reliance was 33,8% which is slightly smaller than was the corresponding average in Beard and Beil's (1994) original study of the game. By leaning on the interpretation of Beard and Beil we can suggest that (i) if a player chooses action A, she relies that her opponent player is rational and chooses also action A and, that (ii) if a player chooses B, she attaches some probability to the non-maximizing behaviour of an opponent player. *Alternatively*, we can interpret that if a player chooses B she attaches value to her own independence, that is, to not having to rely on someone other's choice.<sup>19</sup>

Statistical tests (one-tailed Fisher exact test) brought out some interesting results (see Appendix 3 Tables A3.3-A3.5). An interesting finding was that one's reliance on others' rationality was associated to the choice patterns in the experiment and to the information that was available of the other players. Accordingly, there was an indication that those who relied on the rationality of others, favoured self-oriented choices in all choices until the majority voting procedure was carried out. After this, the 'reliers' shifted towards equity-oriented choices. No such shift happened in the group of 'non-reliers', however. Majority of them favoured the equity-oriented choices in all rounds except the last one (after the group-negotiation).

We can also adopt the interpretation that when a subject chooses the option B in the secure game, she attaches value to her independence of the other subjects' choices. It would then be suggested that maybe the valuing of independence has a relation to the voice dimension in assessing the fairness of the decision mechanism. Fisher's one-tailed exact test (Table 6) shows that this indeed

---

<sup>19</sup> Beard and Beil (1994) refer to Slovic (1987) as a source of the alternative interpretation.

was the case ( $p = .013$ ) in the context of the group-negotiation treatment (but not the other treatments). In other words, when a subject exhibited esteem for independence, she very often *thought she was* an outsider in the group-negotiation procedure.

Table 6. Relying on the rationality of others vs. voice-dimension (group-negotiation treatment)

Relies on the rationality of others	IOS-scale			$p^*$
	Outsider	Insider	Total	
Yes	2	10	12	.013
No	5	1	6	
Total	7	11	18	

The comparison of the behaviour in the dictator and secure game *with* the behaviour in the exclusion game brought some interesting insights into the factors that are in operation in an exclusion game experiment. One way to use dictator game and secure games would be to take them as controlling variables. Another way would be to open with them new interpretations concerning, for example, the factors driving the processes of agreeing to a norm in the exclusion game.

#### 4. Discussion

In general, the results of the exploratory experiment are in line with the results of Sacconi and Faillo's (2005) experiment. In both experiments a significant number of players conformed to the rule they had jointly agreed upon. Since punishing of deviators from the rule was not possible, the adoption of the rule can be interpreted as an activation of the players' inherent valuation of the rule and the general constitutive ideal behind it. Both experiments give lend to the idea, that many people are motivated not only by material consequences of their decisions but also by their willingness to adopt a rule or an ideal. Naturally, the experiment reported here brings only a limited increase to the robustness of the original result by Sacconi and Faillo. In spite of all, it can be suggested that the future investigations should both replicate the result and focus on the procedures of agreeing on a norm to find out what is the best way, in a given situation, to select a norm to get the participants to cooperate for mutual benefit.

The analysis suggests that the future investigation should also pay attention to the fairness issues related to conformist behaviour. An important issue is *the voice condition* with the meaning that (i) some social choice mechanisms are more respectful than others towards the members of the deciding group, and that (ii) some people see a particular process as a fair one (in the sense of reflecting their preferences) while others may think of it as an unfair process external to their own desires preferences. An important tentative finding was that if a player perceived a given mechanism

as a fair way to reach a joint agreement, it was more likely that the player's own choices and her beliefs of others' choices 'coincided' and that the resulting behavioural pattern was conformist.

In addition, the analysis demonstrated that in such a behavioural experiment as the exclusion game experiment is, there are 'other factors' at work, such as learning effects, information effects, other-regarding preferences and the subject's attitude towards others. The 'other factors' may either strengthen or debase the influences of the conformist motivation or fairness impacts and further knowledge of them is important in theory development as well as in designing new experiments. The major 'finding' in the analysis was that the secure game was able to identify players who value their own independence of the actions and decisions of others. This opens new insights into the group dynamics in collective decision making which naturally needs more attention in the future studies.

## **5. Conclusions**

Economists have been interested in fairness of economic allocations and states for quite a long time. Recent theories of social preferences incorporate a handful of new aspects of fairness issues into economic theories. In spite of this, there is still lack of knowledge on how the fairness of procedures affect on the behaviour of individuals. This paper attempts to associate the question of procedural fairness to the study of conformist preferences as an explanation of why individuals in experiments and also in field settings often choose a cooperative or an equity-oriented strategy rather than the strategy which maximizes the personal material payoffs of the acting individual.

The results of the analysis suggest, albeit on the basis of limited evidence, that conformist and process-regarding preferences may outweigh the self-interested preferences in contractarian social dilemma situations. In these situations individual and collective interests are at odds but the individuals have a possibility to jointly agree on a norm to be voluntarily followed later on. In other words, the study demonstrates that norms and fairness can be stronger sources of motivation than the motivation to maximize one's own material outcomes. The operation of reciprocal conformism has already been demonstrated by Sacconi and Faillo (2005), but the present study encourages investigating also the influences of procedural fairness on the behaviour of individuals.

## References

- Aviram, A. (2003) Regulation by Networks. *Brigham Young University Law Review*, 1: 2003.
- Anand, P. (2001) Procedural Fairness in Economic and Social Choice: Evidence from a Survey of Voters, *Journal of Economic Psychology* 22: 247-270.
- Ben-Ner, A. and Putterman, L. (1999) Values and Institutions in Economic Analysis, in Ben-Ner, A. and Putterman, L. (eds.), *Economics, Values, and Organization*, Cambridge University Press.
- Beard, T. and Beil, R. (1994) Do People Rely on the Self-interested Maximization of Others? An Experimental Test, *Management Science*, 40: 252-262.
- Benz, M. (2005) The Relevance of Procedural Utility for Economics, Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 256.
- Bicchieri, C. (2006), *The Grammar of Society: The Nature and Dynamics of Social Norms*, Cambridge University Press.
- Bolton, G. (1991) A Comparative Model of Bargaining: Theory and Evidence, *American Economic Review*, 81: 335-339.
- Bolton, G., Brandts, J. and Axel Ockenfels (2005) Fair Procedures: Evidence from Games Involving Lotteries, *The Economic Journal*, 115: 1054-1076.
- Bolton, G. and Ockenfels, A. (2000) ERC: A Theory of Equity, Reciprocity, and Competition, *American Economic Review*, 90: 166-193.
- Bowles, S. (2004) *Microeconomics: Behavior, Institutions, and Evolution*. Princeton University Press.
- Bruni, L. and Sugden, R. (2007) The Road Not Taken: How Psychology Was Removed From Economics, And How It Might Be Brought Back? *The Economic Journal*, 117: 146-173.
- Charness G., Rabin M., 2002, 'Social Preferences: Some Simple Tests and a New Model', *Quarterly Journal of Economics*, 117: 817-69.
- Cox, J (2004) How to Identify Trust and Reciprocity, *Games and Economic Behavior*, 46: 260-281.
- De Cremer, D. and Stouten, J. (2003) When Do People Find Cooperation Most Justified? The Effect of Trust and Self-Other Merging in Social Dilemmas, *Social Justice Research*, 16: 41-52.
- De Cremer, D., T.R. Tyler and N. den Ouden (2005) Managing Cooperation via Procedural Fairness: The Mediating Influence of Self-Other Merging, *Journal of Economic Psychology*, 26: 393-406.
- Dolan, P., R. Edlin, A. Tsuchiya, and A. Wailoo (2007) It ain't What You Do, It's the Way that You Do It: Characteristics of Procedural Justice and Their Importance in Social Decision-Making, *Journal of Economic Behaviour and Organization*, 64: 157-170
- Faillo, M. and Sacconi, L. (2007) "Norm Compliance: The Contribution of Behavioral Economics Models," Department of Economics Working Papers 0704, Department of Economics, University of Trento, Italia.

Falk, A. and Fischbacher, U. (1998) A Theory of Reciprocity, University of Zurich, IEER working paper.

Fehr E., Schmidt K., 1999, 'A Theory of Fairness, Competition and Cooperation', *Quarterly Journal of Economics*, CXIV, 817-51.

Frey, B, M. Benz, and A Stutzer (2003) Introducing Procedural Utility: Not only What, but also How Matters, *Journal of Institutional and Theoretical Economics*, 160: 377-401.

Geanakopolos, J., D. Pearce and E. Stachetti (1989). "Psychological Games and Sequential Rationality." *Games and Economic Behavior* 1: 60-79.

Gold, N. and Sugden, R. (2007) Theories of Team Agency. *Rationality and Commitment*, in Fabienne Peter and Hans Bernhard Schmid (eds.), Oxford University Press, Oxford.

Grimalda, G. and Sacconi, L. (2005) The Constitution of the Not-For-Profit Organisation: Reciprocal Conformity to Morality, *Constitutional Political Economy*, 16: 249–276.

Guala, F. (2006) Has Game Theory Been Refuted? *Journal of Philosophy*, 103: 139-163.

Hargreaves Heap, S. (1992) Rationality, in Hargreaves Heap, S, Hollis, M, Lyons, B, Sugden, R. and Weale A. (eds.) *The Theory of Choice*. Blackwell.

Hargreaves Heap, S. and Varoufakis, Y. (2004) *Game Theory, a Critical Text*, Routledge, New York.

Lind, E., C. Kulik, M. Ambrose, and M. De Vera Park (1993) Individual and Corporate Dispute Resolution: Using Procedural Fairness as a Decision Heuristic. *Administrative Science Quarterly*, 38: 224-251.

Machina, M. (1989) Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty, *Journal of Economic Literature*, 27: 1622-1668.

Pommerehne, W.W., A. Hart, and F. Schneider (1997) Tragic Choices and Collective Decision-Making: An Empirical Study of Voter Preferences for Alternative Collective Decision-making Mechanisms, *The Economic Journal*, 107: 618-635.

Rabin M., 1993, 'Incorporating Fairness into Game Theory and Economics', *American Economic Review*, 83: 1281-1302.

Rubinstein, A. (1999) Experience from a Course in Game Theory: Pre and Post-class Problem Sets as a Didactic Device", *Games and Economic Behavior*, 28: 155-170

Sacconi, L. (2007) A Social Contract Account for CSR as an Extended Model of Corporate Governance(II): Compliance, Reputation and Reciprocity, *Journal of Business Ethics* 75: 77-96.

Sacconi, L. and Faillo, M. (2005), "Conformity and Reciprocity in the 'Exclusion Game': An Experimental Investigation", University of Trento Economics Working Paper No. 12.

Santos, C. (2006) *The Social Epistemology of Experimental Economics*, Doctoral thesis, Erasmus University of Rotterdam.

Sally, D. (2002) Two Economic Applications of Sympathy, *Journal of Law, Economics, and Organization*, 18: 455-487.

- Sen, A (1997) Maximization and the Act of Choice, *Econometrica*, vol. 65: 745-780.
- Sen, A. (2000) Consequential Evaluation and Practical Reason, *The Journal of Philosophy*, Vol. 97: 477-502.
- Slovick, P. (1974) Perception and Risk, *Science*, 236: 280-285.
- Sugden, R. (1993) The Contractarian Enterprise, in Gauthier, D. and Sugden, R. (eds), *Rationality, Justice and the Social Contract: Themes from Morals by Agreement*. Wheatsheaf: New York.
- Sugden, R. (2000) Team Preferences, *Economics and Philosophy*, 16: 175-204
- Sugden, R. (2005) Experiments as Exhibits and Experiments as Tests, *Journal of Economic Methodology*, 12: 291-302.
- Trautmann, S. (2007) Fehr-Schmidt Process Fairness and Dynamic Consistency. Mimeo. Econometric Institute, Erasmus University Rotterdam.
- Wailoo, A. and Anand, P. (2005) The Nature of procedural Preferences for Health-care Rationing Decisions, *Social Science and Medicine*, 60: 223-236.



**Appendix 1.** Choice tasks in stage one in the experiment

**Question 1**

Imagine, that your professor selects two students in a class. One of the students is assigned the role of player 1 and the other player 2. A sum of money, 10 euros, will be divided between these two players according to player 1's choice in the following clause:

"I will divide 10 euros in the following way: N euros to me and the rest to the other."

The payments are then made anonymously.

**What is your choice of N? Answer \_\_\_\_\_**

**Question 2**

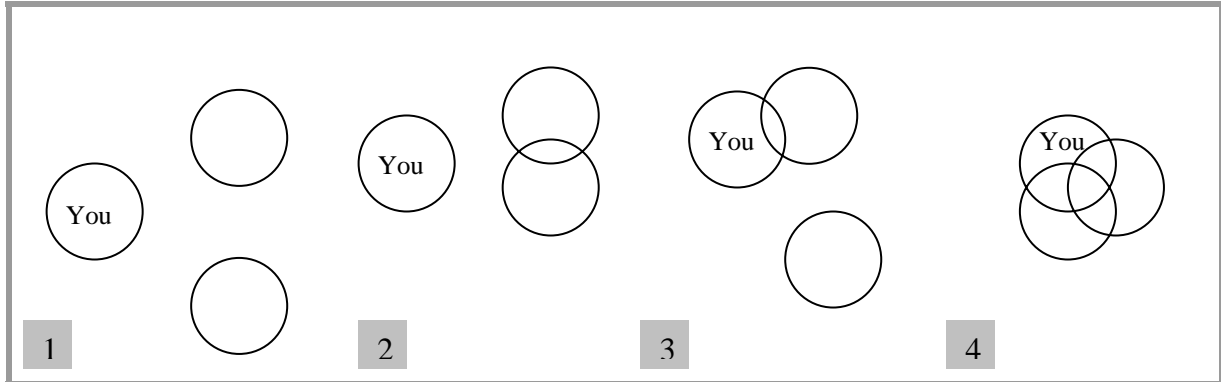
You are player 1 in a two-player game the payoff matrix of which is below (in euros). You choose either A or B without knowing what is player 2's choice (A or B).

		Player 2	
		A	B
Player 1	A	You 5; Other 5	You 0; Other 4
	B	You 2; Other 2	You 3; Other 3

**What is your choice, A or B ? Answer \_\_\_\_\_**

**Appendix 2.** IOS scale task as applied to the group-negotiation

‘Below you see four alternative descriptions of your own standing relative to the other two members of your group in the group-negotiation procedure. Which of the alternatives best describes you own perception?’



**Appendix 3. Tables to section 3.5.**

**Table A3.1.** Rankings of fairness

	Attitude in general				Attitude in the experiment		
	Lottery	Voting	Negotiation	Expert system	Lottery	Voting	Negotiation
Very fair	6	5	4	1	5	3	6
Fair	8	10	11	13	6	14	11
Unfair	3	2	3	3	5	1	1
Very unfair	1	1	-	1	2	-	-

**Table A3.2.** IOS-scale

	Lottery	Voting	Negotiation
No-voice			
Description 1	10	7	3
Description 2	1	1	4
Voice			
Description 3	3	3	1
Description 4	4	7	10

**Table A3.3.** Choices in the secure game against choices in the exclusion game: before treatments

Before lottery				Before voting				Before negotiation			
Choice Relies	E	S	Fisher exact p-value (1-sided)	Choice Relies	E	S	Fisher exact p-value (1-sided)	Choice Relies	E	S	Fisher exact p-value (1-sided)
Yes	5	19	.031	Yes	5	19	.001	Yes	9	15	.09
No	7	5		No	10	2		No	8	4	

**Table A3.4.** Choices in the secure game against choices in the exclusion game: after treatments

After lottery				After voting				After negotiation			
Choice Relies	E	S	Fisher exact p-value (1-sided)	Choice Relies	E	S	Fisher exact p-value (1-sided)	Choice Relies	E	S	Fisher exact p-value (1-sided)
Yes	10	14	.004	Yes	13	10	.417	Yes	13	11	.546
No	11	1		No	8	4		No	6	6	

**Table A3.5.** Choices in the secure game against prediction of other's choices

After lottery				After voting				After negotiation			
Predicts Relies	E	S	Fisher exact p-value (1-sided)	Predicts Relies	E	S	Fisher exact p-value (1-sided)	Predicts Relies	E	S	Fisher exact p-value (1-sided)
Yes	8	16	.006	Yes	21	2	.011	Yes	18	6	.020
No	10	2		No	6	6		No	4	8	