1-1-2010

Inequality Aversion and Efficiency with Ordinal and Cardinal Social Preferences – An Experimental Study

Dorothea K. Herreiner
Loyola Marymount University, dherreiner@lmu.edu

Clemens Puppe
University of Karlsruhe

Recommended Citation

"Inequality Aversion and Efficiency with Ordinal and Cardinal Social Preference – An Experimental Study", *Journal of Economic Behavior and Organization*,76/2, 238-253, November 2010.

Inequality Aversion and Efficiency with Ordinal and Cardinal Social Preferences – An Experimental Study*

DOROTHEA K. HERREINER

Loyola Marymount University
University Hall
1 LMU Drive, Suite 4200
Los Angeles, CA 90045-2659
U.S.A.
dherreiner@lmu.edu

and

CLEMENS PUPPE

Department of Economics
University of Karlsruhe (TH)
D – 76128 Karlsruhe
Germany
puppe@wior.uni-karlsuhe.de

Revised Version, August 20, 2009

* We thank an editor, an associate editor, and two anonymous referees for valuable comments and suggestions. A former version of this paper circulated under the title "Equitable Allocations in Experimental Bargaining Games: Inequality Aversion versus Efficiency" (Bonn Econ Discussion Paper Series, No.29, 2004). We are grateful to Steven Brams, Gary Charness, Heike Henning-Schmidt, Sebastian Kube, Karl Schlag, Reinhard Selten, Avner Shaked and William Thomson for discussions and valuable comments on earlier versions of this and our related work. We thank Sebastian Kube and Javier Sanchez Monzon for research assistance, and Thorsten Chmura, Thomas Pitz and the staff at the Experimental Lab at the University of Bonn for helping to conduct the experiments. Financial support by the Department of Economics at the University of Bonn and the Fletcher Family Research Grant at Bowdoin College is also gratefully acknowledged.

Abstract In this paper, we report on a series of free-form bargaining experiments in which two players have to distribute four indivisible goods among themselves. In one treatment, players are informed about the monetary payoffs associated with each bundle of goods; in a second treatment only the ordinal ranking of the bundles is given. We find that in both cases, inequality aversion plays a prominent role. In the ordinal treatment, individuals apparently use the ranks in the respective preference orderings over bundles of goods as a substitute for the unknown monetary value. Allocations that distribute the value (money or ranks, respectively) most equally serve as natural "reference points" for the bargaining processes. Frequently, such "equal split" allocations are chosen by our subjects even though they are Pareto dominated; but also if they are rejected for that reason they matter in a specific way: whether a Pareto optimal allocation is chosen or not depends on whether or not it is a Pareto improvement *relative to* the equal split. Interestingly, we find much less Pareto-damaging behavior due to inequality aversion in the ordinal treatment.

1. Introduction

Recent research in explaining observed behavior of individuals in laboratory experiments has focused on the question of how to model the agents' *distributional preferences*, see, e.g., Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2002) (henceforth F&S, B&O and C&R, respectively). The common assumption in these models is that agents are not only motivated by their own material payoff but by the entire ditribution of monetary rewards. Specifically, F&S and B&O suggest parametric forms of the utility function incorporating different notions of *inequality aversion* according to which utility decreases with the differences in individual payoffs. By contrast, C&R propose a model of *social-welfare* preferences according to which agents are concerned with maximizing a combination of the aggregate payoff for the group and the payoff of the worst-off individual. The two approaches have been compared and tested against each other by Engelmann and Strobel (2004).[1]

By assigning significance to differences and sums of monetary rewards, the proposed models of social preferences use individual utility information in a *cardinal* and *interpersonally comparable* way. While this can be justified, e.g. by assuming quasi-linearity of the underlying preferences, it also shows that the applicability of the existing models is restricted to situations in which individual monetary rewards are known to all agents and in which preferences over allocations can be adequately described in terms of the distribution of monetary rewards.

The purpose of the present paper is to demonstrate that the basic intuitions behind the distributional preference approach can be fruitfully applied in more general situations. To this end, we conducted a series of free-form bargaining experiments in which two players had to jointly determine an allocation of four indivisible goods. In one treatment both agents were informed about the specific monetary value associated with the bundles of goods for each player (the same bundle usually had different monetary value for the two players). In the other treatment, each player was only informed about her own and the opponent's ordinal ranking of the bundles, i.e. only the ordinal ranking of the monetary payments associated with each bundle was given. Despite the lack of numerical

[1] See, among others, also the comments by Bolton and Ockenfels (2006) and Fehr, Naef and Schmidt (2006), as well as the reply by Engelmann and Strobel (2006).

payoff information in the latter treatment, we find that individuals rely on interpersonal comparisons also in this case. Indeed, we find strong evidence that agents use the rank of a bundle in the respective preference ordering as a substitute for its unknown monetary value. Taking these ranks as the basis for interpersonal comparisons, the motives behind the formation of distributional preferences, such as inequality aversion or social concerns in general, are relevant also in the treatment with ordinal information. In fact, the comparison between the two treatments reveals that individual behavior can be accounted for by a simple unifying *qualitative* theory of distributional preferences. Specifically, the outcomes that we observed in our bargaining experiments suggest that a significant proportion of agents' behavior is guided by the following rule:

> Conditional Pareto Improvement from Equal Split (CPIES):
> First, determine the most equal distribution of rewards. If this allocation is Pareto optimal, then choose it. Otherwise, if there is the possibility to make everyone better off, implement such a Pareto improvement *provided* that this does not create "too much" inequality.

If the monetary rewards are known, the "most equal" distributions are of course the ones with minimal difference of the numerical payoffs for the two agents.[2] If, on the other hand, only the ordinal rankings of the bundles of goods are given, then the "most equal" distributions are those with minimal difference of the ranks in the respective preference orderings. Similarly, "too much inequality" is to be understood in terms of differences in monetary payoffs and ranks, respectively. Of course, how much precisely "too much" is, depends on individual preferences and varies from subject to subject.

The above rule combines elements of the inequality aversion approach of F&S and B&O on the one hand, and the social-welfare preference approach of C&R on the other. With the former it shares the important role played by interpersonal equality, with

[2] In our setting with two players, there are at most two such distributions in the feasible set. With more than two players, the meaning of "most equal" distribution can be made precise using the theory of inequality measurement; for instance, by applying the (partial) criterion of Lorenz dominance.

the latter the demand for Pareto optimality (in the payoff space).[3] Interpersonal inequality plays a twofold role here. First, the absence of inequality determines an initial *reference* point for the bargaining problem. Secondly, it serves as a *constraint* in the process of achieving a Pareto optimal outcome. In contrast to C&R's results, we systematically find Pareto-damaging behavior in the treatment with known monetary rewards.[4] Interestingly, however, such behavior is only very rarely observed in the ordinal treatment. Our conjecture is that this is due to the uncertainty about the differences in final payments associated with differences in ordinal ranks. Indeed, it seems that rank inequality becomes acceptable because it does not *necessarily* correspond to unequal monetary payoffs. One conclusion from our study is thus that, by making inequality precisely quantifiable, monetary payoff information hinders the realization of Pareto improvements.

The CPIES rule admits two different interpretations. The first is purely outcome-oriented: whether or not an allocation is compatible with the CPIES rule can be decided simply by looking at the resulting inequality and by determining whether or not it is a Pareto improvement relative to the most equal allocation. The second interpretation of the CPIES rule is as a proper procedure according to which bargaining partners first determine a "disagreement point" which then serves as the reference distribution for the later bargaining process. In Section 4 below, we look at both interpretations. In terms of statistical analysis, the relevance of the CPIES rule is more easily tested in its outcome-oriented version. On the other hand, an analysis of the communication protocols of the experiments shows that the CPIES rule indeed frequently materializes in the procedural sense: the equal reference distribution is proposed, or mentioned in the discussion although not necessarily proposed, and then the bargaining partners either settle on a Pareto improvement from there, or choose the equal distribution.

Our experimental design differs from the literature in several respects. First, while most of the existing studies on distributional preferences have focused on variants of

[3] Pareto optimality is defined in *payoff space* as opposed to utility space, since the relevant notion of optimality here is based on the distribution of material payoffs (respectively, ranks), not on the subjective distributional preferences.

[4] C&R's model does allow for inequality aversion and is capable to incorporate Pareto-damaging behavior for suitable values of the parameters. However, C&R's experimental results suggest parameter values that exclude Pareto-damaging behavior.

either dictator or ultimatum games, we consider free-form bargaining here. In terms of experimental set-up, we thus follow the literature initiated by Roth and Malouf (1979) who tested the predictive power of the Nash bargaining solution using an unstructured bargaining process (see also the subsequent literature, in particular Roth and Murnighan (1982), and the review in Kagel and Roth (1995)). The main reason for the departure of our experimental design from the existing literature on social preferences was the conjectured presence of a procedural aspect influencing the outcome. For instance, the role of the equal distribution as a reference point is much easier uncovered in an unstructured bargaining process than in a rigid strategic game. Qualitatively, our results confirm findings of the early study by Kalisch et al. (1953). Using an informal bargaining process, these authors tested solution concepts of cooperative game theory, and in particular, how members of a coalition would share the joint surplus among each other. One of their findings was that in the process of coalition formation, the “core” or founding members (mostly two individuals) often split the initial surplus equally. While our context is clearly different, the role of the equal distribution as reference point for the bargaining process emerges here in a pronounced way as well.[5]

Data generated by free-form bargaining processes are arguably more difficult to interpret than data from more structured bargaining problems such as dictator and ultimatum games, or alternate offer games. But our experimental design also has evident advantages in terms of exploration of the motives underlying bargaining behavior.[6] In particular, data from free-form bargaining are much better suited to contribute to our understanding of the procedural and cognitive aspects involved. The growing but still small literature on procedural justice suggests that the way in which final resource allocations are brought about has a significant impact on their acceptability. The role that procedures play in resource allocation problems has been investigated from different perspectives. For instance, Bolton, Brandts and Ockenfels (2005) find that unequal distributions are more easily accepted if they are the outcome of an *ex-ante* fair procedure, say of a fair lottery. Shor (2007), on the other hand, examines the effects of the distribution of decision power among individuals. Our study complements this

[5] The role of reference points in ultimatum bargaining has recently been emphasized by Bolton and Ockenfels (2005).
[6] Cf. Camerer (2003, Ch.4).

literature by examining how the bargaining proposals evolve over time and by directly looking at our subjects' arguments and reasoning during the bargaining process in order to identify regularities and recurring patterns. [7] We view the present paper as a first explorative step and hope that it will stimulate further research in this direction.

Another distinctive feature of our experimental design is the framing of the decision problem as one of distributing indivisible goods. Hence, the feasible payoff distributions are explicitly derived from an underlying economic allocation.[8] In contrast to Kalisch et al. (1953), Roth and Malouf (1979), or Shor (2007) the set of feasible utility distributions is thus restricted.[9] More importantly, our design allows us to induce purely ordinal preference information and to compare the corresponding results with those obtained under full (cardinal) payoff information.

The remainder of the paper is organized as follows. Section 2 describes the experimental design and Section 3 the division problems that we tested. Section 4 presents the results, treating the outcome-oriented and process-oriented interpretation of the CPIES rule in two separate subsections. Section 5 concludes.

2. Experimental Design

We ran four different bargaining experiments (EXP I – IV) in six sessions each. Eight subjects participated in each session, so that the total number of subjects was 192 (=8*6*4). Each session consisted of five different rounds (R1 – R5). Subjects were re-matched in pairs after each round; no subjects met twice.[10] We treated the results in each round as independent observations.[11] The experiments took place in June and July 2001

---

[7] There is a large related literature on procedural fairness in psychology, see, e.g., Lind and Tyler (1988) for a rich monograph, and Tyler and Lind (2000) for a more recent survey.

[8] We employed a similar design in Herreiner and Puppe (2004). Part of the data reported here were also used in that paper in order to test for the relevance of the criterion of envy-freeness.

[9] Another minor difference to Roth and Malouf (1979) is that proposals made by bargaining partners could be withdrawn and became binding only after confirmation by both partners.

[10] For the matching of subjects, see Appendix 1.

[11] Although subjects never met twice, it is obviously possible that the behavior and/or experience in earlier rounds had an impact on how subjects behaved in later rounds. We tested whether behavior in comparable or even identical problems was the same irrespective of what happened before. Specifically, we compared EXP I R4, EXP II R1, EXP II R5, and also EXP I R5, EXP II R2, EXP II R4; lastly, we compared the identical problems EXP I R3 and EXP II R3. The null hypothesis that the behavior is the same in these

(EXP I, II and III) and November 2002 (EXP IV) at the University of Bonn. Almost all subjects were students at the University of Bonn, most of them in economics or law.

In each round of each experiment, the subject pairs had to bargain over the distribution of four indivisible goods, denoted by A,B,C and D. Bargaining partners had to reach an agreement within[12] 10 minutes, otherwise neither received anything. Partners communicated via computer by sending proposals at any time. A proposal consisted of a specification of a distribution of the four goods between the two players. Each good could be given to only one player, and all four goods had to be distributed (no free disposal). In addition to sending proposals, bargaining partners had the possibility to support their proposals by verbal messages that they sent via a "chat" device on the screen (see the instructions in Appendix 4 for screen examples). When two partners had proposed the same allocation of goods, they were asked for confirmation. If both confirmed, the goods were allocated accordingly and the round was over for them. If two partners could not reach an agreement within 10 minutes neither player received a good and the round was over.

Each bundle of goods corresponded to a value in experimental currency ("Taler"). The Taler values of the bundles ranged from 0 (for the empty bundle) to 100 (for all four goods combined). The conversion rate was 24 Taler for 1 €.[13] Importantly, the same bundle typically had a different value for different players. Moreover, the value functions were not additive, i.e. the value of a combination of goods was in general different from the sum of the values of its components. Thus, the valuations allowed for complementarities between goods. For instance, in Table 1 below, good B is worth 7 Taler and good C is worth 3 Taler for player 2; however, the bundle BC that combines them is worth 75 Taler. Also, the bundle BC is more valuable than the bundle AC

---

cases cannot be rejected at any relevant significance level (weaker than 10%) based on two tailed exact Fisher tests for any pairwise comparison of the relevant rounds across experiments. The Fisher tests considered the two most frequently chosen allocations for each case. χ^2 tests based on all chosen allocations yielded similar results. It thus seems justified to treat the results in each round as independent observations (see also C&R, p.827 for a similar approach).
The regression results in Appendix 3 support the independence assumption.

[12] Once all pairs had agreed on an allocation, the round was over for everybody and the next round started immediately.

[13] In the experiments before January 1, 2002, the conversion rate was 12 Taler for 1 DM.

although good A taken in isolation is more valuable (8 Taler) than good B taken in isolation.[14]

There were two different treatments. In treatment *CARD* each subject was informed about the precise Taler value of each bundle for either player (as in Table 1 below). In treatment *ORD* the subjects were only informed about the ordinal ranking of the bundles of goods for either player and about the fact that the empty bundle earned 0 whereas all four goods combined yielded 100 Taler (see Table 2 below).[15] Note that complementarities between goods can also be present in the ordinal treatment. For

Player 1			Player 2	
Monetary Payoffs	Bundles		Bundles	Monetary Payoffs
100	ABCD		ABCD	100
98	ABC		ABC	97
95	ABD		ABD	96
93	CBD		ACD	91
83	ACD		CBD	88
66	AB		BC	75
57	CD		AC	45
53	BC		BD	42
46	AD		CD	40
45	BD		AB	28
20	AC		AD	19
9	B		A	8
5	A		B	7
3	C		C	3
1	D		D	2
0	-		-	0

Table 1: Example of treatment CARD (EXP I, R3 = EXP II, R3)

Player 1		Player 2
Bundles		Bundles
ABCD		ABCD
BCD		ABC
ABD		BCD
ABC		ABD
ACD		ACD
AC		AB
AD		AC
AB		CD
BC		BC
CD		BD
BD		AD
D		B
C		C
A		A
B		D
-		-

Table 2: Ex. of treatment ORD (EXP III, R3)

[14] The complementarities between goods forced subjects to focus on the entire ranking of bundles of goods, which is important for the interpretation of our results.

[15] During the instructions, the subjects were given several (extreme) examples in order to show that the unknown cardinal numbers behind the ordinal rankings could be very different, in particular far away from an equidistant scale.

instance, B and C are the two most valuable single goods for player 2 in Table 2; nevertheless, the bundle BC is not the most valuable combination of two goods. Experiments EXP I and II involved the *CARD* and *ORD* treatments, and EXP III and IV only the *ORD* treatment.

3. The Bargaining Problems

We now describe the bargaining problems that we tested. Each experiment is summarized in a separate table that shows the data of the examples used in each of the

R1				R2				R3				R4				R5			
n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.		P1	P2			P1	P2			P1	P2	
100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100
95	ABC	BCD	98	98	ABC	BCD	94	98	ABC	ABC	97	95	ABC	ABD	98	98	ABC	BCD	94
92	ACD	ABD	95	96	ACD	ABD	90	95	ABD	ABD	96	92	BCD	ACD	95	96	ACD	ABD	90
89	BCD	ABC	87	92	BCD	ACD	86	93	BCD	ACD	91	89	ABD	ABC	87	92	BCD	ABC	86
82	ABD	ACD	84	88	ABD	ABC	81	83	ACD	BCD	88	82	ACD	BCD	84	88	ABD	ACD	81
60	BC	BD	64	60	BD	CD	64	66	AB	BC	75	60	AB	AD	64	60	BC	BD	64
55	AB	BC	47	45	AC	BC	53	57	CD	AC	45	55	AC	AB	47	45	AB	BC	53
50	CD	AC	43	40	CD	AD	50	53	BC	BD	42	50	BD	BC	43	40	CD	AC	50
46	AD	CD	38	36	AB	AC	44	46	AD	CD	40	46	CD	BD	38	36	AD	CD	44
35	BD	AB	30	30	AD	BD	32	45	BD	AB	28	35	AD	AC	30	30	BD	AB	32
28	AC	AD	27	28	BC	AB	26	20	AC	AD	19	28	BC	CD	27	28	AC	AD	26
15	C	B	17	9	C	D	19	9	B	A	8	15	B	A	17	9	C	B	19
12	A	D	11	8	A	B	15	5	A	B	7	12	C	D	11	8	A	D	15
7	B	C	5	5	B	C	10	3	C	C	3	7	A	B	5	5	B	C	10
5	D	A	4	2	D	A	7	1	D	D	2	5	D	C	4	2	D	A	7
0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0

P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ
AB	**CD**	10	PO	2	**AB**	**CD**	10	PO	3	**BD**	**AC**	14		0	**CD**	**AB**	21	PO	1	**AB**	**CD**	21	PO	1
AD	**BC**	6	PO	2	**BD**	**AC**	8	PO	3	**AD**	**BC**	7	PO	29	**AC**	**BD**	1	PO	17	**CD**	**AB**	1		8
BD	**AC**	2		2	**AD**	**BC**	2		3	**AB**	**CD**	2	PO	26						**C**	**ABD**	1	PO	81
AC	**BD**	2	PO	5	**BC**	**AD**	1		3															
CD	**AB**	1		2	**AC**	**BD**	1		3															
ACD	**B**	1	PO	9																				
ABCD	**-**	1	PO	15																				
		23					22					23					22					23		

Table 3: EXP I

The upper part of the table shows the two rankings participants (P1, P2) saw during the 5 rounds of the experiment (R1-R5); "n.a." indicates an ordinal treatment (R1, R2), where players did not see the monetary payoffs (in light grey). In the cardinal treatments (R3-R5), players saw the ranking of bundles with corresponding monetary payoffs next to them, as shown above.
The lower part of the table lists the allocations (P1's and P2's bundle) that were chosen from most to least frequent (#), identifying whether an allocation was Pareto optimal; the last column (Δ) shows the rank differences in the ordinal treatment, respectively the monetary payoff differences in the cardinal treatment.

five rounds in the upper part (rows 2 – 18) and summarizes the results in the lower part (rows 20 – 27).

Table 3 above refers to EXP I. In row 2, "n.a." stands for "not available" and refers to the fact that in the *ORD* treatment the later Taler payments associated with the bundles of goods (printed in light gray in the tables) were not known by the two players at the time of bargaining. In the upper part of the table, the columns P1 and P2 indicate the ordinal rankings of the bundles for player 1 and 2, respectively. The lower part of the table (rows 20 – 27) shows the chosen distributions of goods (column P1 indicates player 1's bundle and column P2 player 2's bundle). The frequency of the respective choices is shown in column 3 (#). Thus, for instance, the allocation (AB,CD) (i.e. the bundle AB for player 1, and the bundle CD for player 2) was chosen by 10 bargaining partners in R1 of EXP I. The fourth column of the bottom part indicates whether the respective distribution is Pareto optimal. Finally, the fifth column (Δ) gives the difference in the ranks of the respective bundles in the preference orderings of the two players, and, if applicable, also the difference in Taler payoffs. Thus, for instance in R3 of EXP I, the Pareto optimal distribution (AD,BC) was chosen seven times, it involves a difference of 3 ranks, and a payoff difference of 29 Taler.

EXP I involved both the *ORD* and the *CARD* treatments, as the first two rounds were without numerical payoff information. However, these two rounds were designed for a different purpose and are evaluated in Herreiner and Puppe (2004); the same applies to the example in R5 in EXP III (see Table 5 below).[16] The distinctive feature of the division problem in R3 of EXP I is the tension between equality and efficiency. The distribution (BD,AC) of goods results in the equal distribution (45,45) of monetary payoffs; however, this distribution is strictly dominated by the distribution (AD,BC) that yields payoffs (46,75). By contrast, the most equal distributions in the last two rounds of

[16] Specifically, these examples were designed to test the criterion of envy freeness. For instance, in R1 of EXP I the distribution (AB,CD) is *envy free* since each player prefers her own bundle to that of her partner. This property does not hold, say, for the distribution (AD,BC), where player 1 would rather have her partner's bundle. In R1 and R2 of EXP I and R5 of EXP III there is no unambiguous distribution which may serve as reference point for the CPIES procedure; by consequence, the observed choices (see Tables 3 and 5) support the CPIES prediction in a trivial way. The results of all three examples do confirm the other findings reported here, in particular the result that, in contrast to the cardinal treatment, Pareto optimality is often achieved in the ordinal treatment.

EXP I, i.e. the distribution (CD,AB) in R4 and the distribution (AB,CD) in R5, are both Pareto optimal.

Table 4 below shows the examples and results of EXP II which involved only the *CARD* treatment. The example tested in R3 with the trade-off between equality and efficiency is identical to R3 of EXP I. In all other rounds, the unambiguously most equal distribution is Pareto optimal.

R1				R2				R3				R4				R5			
	P1	P2			P1	P2			P1	P2			P1	P2			P1	P2	
100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100
95	ABC	BCD	98	98	ABC	BCD	94	98	ABC	ABC	97	97	ABC	BCD	95	96	ABC	BCD	97
92	ACD	ABD	95	96	ACD	ABD	90	95	ABD	ABD	96	95	ACD	ABD	91	91	ACD	ABD	93
89	BCD	ABC	87	92	BCD	ACD	86	93	BCD	ACD	91	93	BCD	ABC	86	90	BCD	ABC	88
82	ABD	ACD	84	88	ABD	ABC	81	83	ACD	BCD	88	87	ABD	ACD	82	83	ABD	ACD	86
60	BC	BD	64	60	BC	BD	64	66	AB	BC	75	60	BC	BD	64	60	BC	BD	64
55	AB	BC	47	45	AB	BC	53	57	CD	AC	45	47	AB	BC	52	56	AB	BC	46
50	CD	AC	43	40	CD	AC	50	53	BC	BD	42	42	CD	AC	51	52	CD	AC	41
46	AD	CD	38	36	AD	CD	44	46	AD	CD	40	35	AD	CD	46	45	AD	CD	39
35	BD	AB	30	30	BD	AB	32	45	BD	AB	28	33	BD	AB	32	39	BD	AB	35
28	AC	AD	27	28	AC	AD	26	20	AC	AD	19	29	AC	AD	28	31	AC	AD	30
15	C	B	17	9	C	D	19	9	B	A	8	9	C	B	18	14	C	B	16
12	A	D	11	8	A	B	15	5	A	B	7	7	A	D	17	13	A	D	14
7	B	C	5	5	B	C	10	3	C	C	3	6	B	C	11	8	B	C	7
5	D	A	4	2	D	A	7	1	D	D	2	3	D	A	6	2	D	A	4
0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0

P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ
AD	**BC**	19	PO	1	**AB**	**CD**	22	PO	1	**AD**	**BC**	12	PO	29	**AB**	**CD**	23	PO	1	**AD**	**BC**	22	PO	1
AB	**CD**	2	PO	17	**BC**	**AD**	1	PO	34	**BD**	**AC**	8		0	**CD**	**AB**	1		10	**BC**	**AD**	1	PO	30
ACD	**B**	1	PO	75																**AB**	**CD**	1	PO	17
BD	**AC**	1		8																				
CD	**AB**	1		20																				
		24					23					20					24					24		

Table 4: EXP II

The upper part of the table shows the two rankings participants (P1, P2) saw during the 5 rounds of the experiment (R1-R5). Players saw the ranking of bundles with corresponding monetary payoffs next to them, as shown above.
The lower part of the table lists the allocations (P1's and P2's bundle) that were chosen from most to least frequent (#), identifying whether an allocation was Pareto optimal, and indicating the corresponding monetary payoff differences (Δ).

EXP III and IV only involved the *ORD* treatment. Table 5 summarizes EXP III. R1 in EXP III was designed to test the impact of the number of goods on the choice of allocations. Note that the distributions (AB,CD) and (ABC,D) are both Pareto improvements from the reference distribution (AD,BC), and both involve a rank difference of 3. Nevertheless, (AB,CD) is chosen much more frequently (14 times versus

R1				R2				R3				R4				R5			
n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.
100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100
95	ACD	BCD	90	98	ABC	ACD	94	98	BCD	ABC	97	85	BCD	ACD	85	98	ABC	BCD	94
80	BCD	ACD	80	96	ABD	ABD	90	95	ABD	BCD	96	80	ACD	BCD	80	96	ACD	ABD	90
70	CD	ABD	70	92	BCD	BCD	86	93	ABC	ABD	91	60	CD	ABD	73	92	BCD	ACD	86
65	ABC	CD	48	88	ACD	ABC	81	83	ACD	ACD	88	40	ABC	CD	67	88	ABD	ABC	81
63	ABD	BD	45	60	AC	BC	58	66	AC	AB	75	38	ABD	AD	62	60	BD	CD	64
50	BD	AD	40	45	BC	AC	53	57	AD	AC	45	32	AD	BD	57	45	AC	BC	53
32	AB	D	35	38	AD	AB	50	53	AB	CD	42	24	BC	D	50	40	CD	AD	50
30	AC	ABC	32	36	AB	CD	44	46	BC	BC	40	23	AB	ABC	49	36	AB	AC	44
25	AD	BC	30	30	BD	AD	32	45	CD	BD	28	20	BD	AC	44	30	AD	BD	32
20	BC	AB	28	28	CD	BD	26	20	BD	AD	19	19	AC	AB	37	28	BC	AB	26
15	B	AC	20	9	B	D	19	9	D	B	8	9	A	BC	30	9	C	D	19
10	D	C	15	8	A	A	15	5	C	C	7	7	D	C	9	8	A	B	15
5	C	A	10	5	C	C	10	3	A	A	3	5	C	B	8	5	B	C	10
2	A	B	5	2	D	B	7	1	B	D	2	2	B	A	1	2	D	A	7
0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0

P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ	P1	P2	#		Δ
AB	**CD**	14	PO	3	**AD**	**BC**	17	PO	2	**AB**	**CD**	19	PO	0	**BC**	**AD**	13	PO	2	**BD**	**AC**	11	PO	3
ABC	**D**	3	PO	3	**AC**	**BD**	4	PO	5	**AD**	**BC**	1	PO	2	**ABC**	**D**	7	PO	3	**AB**	**CD**	7	PO	3
AC	**BD**	3		3	**B**	**ACD**	1	PO	10	**CD**	**AB**	3	PO	4	**AB**	**CD**	1	PO	4	**BC**	**AD**	1		3
CD	**AB**	2	PO	7	**AB**	**CD**	2		0						**AC**	**BD**	1		4	**ABC**	**D**	1	PO	10
															ABD	**C**	1		7					
		22					24					23					23					20		

Table 5: EXP III

The upper part of the table shows the two rankings participants (P1, P2) saw during the 5 rounds of the experiment (R1-R5), where "n.a." indicates an ordinal treatment in which players did not see the monetary payoffs (in light grey).
The lower part of the table lists the allocations (P1's and P2's bundle) that were chosen from most to least frequent (#), identifying whether an allocation was Pareto optimal, and indicating the corresponding rank differences (Δ).

3 times).[17] We conjecture that this is due to the more equal *number* of goods distributed. R2 and R3 illustrate the CPIES procedure: in R2 the most frequently chosen distribution (AD,BC) is the unique Pareto improvement relative to the reference distribution (AB,CD). In R3, the reference distribution (AB,CD) itself is Pareto optimal. In R4, there are several Pareto improvements relative to the reference distribution (BD,AC) that was never chosen. Among these, the most frequently chosen allocation (BC,AD) is the unique distribution that minimizes the rank difference, whereas the second-most chosen allocation (ABC,D) is a lexicographic refinement of the "Rawlsian" solution (first minimize the rank of the worst off player; if this is not unique, minimize among all such distributions the rank of the better off player).[18] In R5 there is no unambiguous reference distribution since the minimal rank difference of 3 is achieved by all distributions involving two goods for each player. The CPIES procedure is thus not applicable here. Nevertheless, also this example confirms the prediction of a rank-difference-minimizing distribution among all Pareto optimal distributions (chosen in 90% = 18/20 of all cases).

Table 6 summarizes EXP IV. The distinctive feature of the example in R1 is that there are several Pareto improvements from the reference distribution (AD,BC), not all of which are Pareto optimal. Indeed, the rank-difference-minimizing Pareto improvement (AC,BD) is itself not Pareto optimal. As the examples R3 of EXP I and EXP II, also this example illustrates a tension between inequality and efficiency since 26% (= 6/23) of the observed agreements are not Pareto optimal. Interestingly, the effect is less pronounced here than in the cardinal treatment with full payoff information (see Section 4 for further discussion). The results in R2 very clearly illustrate the role of the rank difference, since the feasible Pareto improvements (AC,BD) and (CD,AB) from the reference bundle (AB,CD) are never chosen, whereas the Pareto improvement (BC,AD) with rank difference 1 is chosen – a rank difference of 3 is apparently considered too large. In addition to R4 of EXP III, R3 in EXP IV is the only example where a rank difference minimizing Pareto optimal distribution does not automatically coincide with the

[17] The null hypothesis that both Pareto improvements are chosen equally likely can be rejected at p=0.006 based on a binomial test.

[18] Note also that, in contrast to the distribution (BC,AD), the distribution (ABC,D) is envy free. In this example, there are thus several conflicting criteria of quite heterogeneous nature.

R1				R2				R3				R4				R5			
n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.	n.a.	P1	P2	n.a.
100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100	100	ABCD	ABCD	100
98	BCD	ACD	87	92	BCD	ABD	97	95	ABC	ABC	85	98	BCD	ABD	87	92	BCD	ABD	97
92	ABC	ABD	77	87	ACD	ABC	87	90	ACD	BCD	75	92	ABC	ACD	77	87	ABC	ABC	87
86	ABD	BCD	62	81	ABC	BCD	72	83	ABD	ABD	66	86	ACD	BCD	62	81	ACD	ACD	72
77	ACD	ABC	60	70	ABD	ACD	70	77	BCD	ACD	62	77	ABD	ABC	60	70	ABD	BCD	70
60	BC	CD	45	63	CD	BD	49	56	AB	AD	51	60	BC	AB	45	63	CD	BD	49
55	BD	AB	43	53	BD	AD	43	50	AC	BC	46	55	BD	BD	43	53	BD	AD	43
42	AB	BD	40	49	BC	BC	40	46	BC	AC	42	42	CD	BC	40	49	BC	AC	40
39	AC	AD	39	42	AC	AB	36	39	AD	BD	35	39	AB	AD	39	42	AC	CD	36
35	AD	BC	36	35	AB	CD	32	32	**CD**	CD	31	35	AC	CD	36	35	AB	AB	32
31	CD	AC	33	30	AD	AC	29	24	BD	AB	28	31	AD	AC	33	30	AD	BC	29
12	B	D	24	16	C	B	24	19	A	B	27	12	C	D	24	16	C	A	24
9	C	A	15	12	D	D	19	13	C	C	17	9	B	A	15	12	B	B	19
3	A	C	7	9	B	A	15	9	B	A	10	3	D	B	7	9	D	D	15
2	D	B	2	8	A	C	8	5	D	D	7	2	A	C	2	8	A	C	8
0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-	0

Round	P1	P2	#		Δ
R1	**AB**	**CD**	14	PO	2
R1	**AC**	**BD**	4		1
R1	**BC**	**AD**	4	PO	3
R1	**AD**	**BC**	1		0
R1			23		
R2	**BC**	**AD**	21	PO	1
R2	**BD**	**AC**	2		4
R2	**AB**	**CD**	1		0
R2			24		
R3	**BC**	**AD**	18	PO	2
R3	**AC**	**BD**	4	PO	2
R3			22		
R4	**CD**	**AB**	22	PO	2
R4	**BC**	**AD**	1	PO	3
R4	**AB**	**CD**	1		1
R4			24		
R5	**BD**	**AC**	14	PO	1
R5	**BC**	**AD**	8	PO	1
R5	**AC**	**BD**	2	PO	3
R5			24		

Table 6: EXP IV

The upper part of the table shows the two rankings participants (P1, P2) saw during the 5 rounds of the experiment (R1-R5), where "n.a." indicates an ordinal treatment in which players did not see the monetary payoffs (in light grey).
The lower part of the table lists the allocations (P1's and P2's bundle) that were chosen from most to least frequent (#), identifying whether an allocation was Pareto optimal, and indicating the corresponding rank differences (Δ).

"Rawlsian" solution, i.e. the maximin solution that minimizes the rank of the worst off player. The allocations (AC,BD) and (BC,AD) have the same rank difference of 2 and are both Pareto optimal, but only (BC,AD) minimizes the rank of the worst off player, and consequently also minimizes the rank sum. Remarkably, it is chosen much more frequently, namely in 82% (= 18/22) of all cases.[19] The two final rounds R4 and R5 once again confirm the role of the three aspects of rank difference, Pareto optimality and

[19] Note that, compared to (AC,BD), the Rawlsian distribution (BC,AD) not only has a smaller sum of ranks for the two players but is also envy free. Therefore, it is not evident which criterion is responsible for the result.

Pareto superiority relative to the reference distribution. Note that in R5, just as in R1 of EXP IV, there is a Pareto improvement from the equal split distribution (AD,BC) that is itself not Pareto optimal, namely the distribution (AB,CD) which is dominated by all of the chosen distributions; in contrast to the allocation (AC,BD) in R1 – the only Pareto improvement with a rank difference of 1 – the distribution (AB,CD) is never chosen in R5, where there exist other Pareto improvements with the same rank difference of 1.

4. Results

The average earning over all experiments and all subjects was about € 9 (in approximately one hour).[20] In fewer than 5% of our observations (23 of 480), bargaining partners did not reach an agreement and thus earned zero in the respective round. The majority of rounds lasted 9 minutes or longer since some subject pairs bargained until shortly before the deadline. However, many agreements were reached much quicker. For instance, in EXP II the median bargaining time ranged from 130 seconds in R5 to 340 seconds in R3.

The following analysis is divided in two subsections. First, we present the results in terms of the final allocations on which bargaining partners settled. In the second part, we then study selected bargaining problems also from a procedural perspective and look at the way in which proposals evolved over time.

4.1 Results in Terms of Final Allocations

The following table summarizes the results in treatment *CARD*. Each entry in the table gives the percentage of the chosen final allocations in the corresponding round (column) that satisfy the criterion corresponding to the row. The first row quantifies how many distributions were Pareto optimal. The second row gives the share of chosen distributions with minimal difference in numerical payoffs (unique in each case), while the third row quantifies how many distributions were consistent with the CPIES procedure. Finally, the

[20] Average earnings were thus comparable to the usual wage for student jobs at the University of Bonn.

last row indicates the total number of agreements reached in the respective round (maximally 24).

	EXP I			EXP II					Total
	R3	R4	R5	R1	R2	R3	R4	R5	
Pareto Optimal	39%	100%	96%	92%	100%	60%	96%	100%	86%
Min Diff ("Reference Point")	61%	95%	91%	79%	96%	40%	96%	92%	82%
CPIES	91%	95%	91%	79%	96%	100%	96%	92%	92%
Number of Agreements	23	22	23	24	23	20	24	24	183

Table 7: Results in treatment CARD (EXP I & EXP II)

The table quantifies the frequency of final allocations in the different bargaining problems compatible with the criteria given in the rows.

Of particular interest are R3 in EXP I and R3 in EXP II in which the same division problem was given. Here "equal split" involves the allocation (BD,AC) with a resulting payoff distribution of (45,45), cf. Tables 3 and 4. In R3 of EXP I, 61% of subject pairs choose this allocation, and 40% in R3 of EXP II. The interesting fact is that this allocation is not Pareto optimal; the allocation (AD,BC) Pareto dominates the equal split with a resulting payoff distribution of (46,75). Note that therefore also both the "Rawlsian" maximin criterion and the utilitarian criterion ("maximize the sum of the individual payoffs") favor (AD,BC) over the equal split. Thus, since a significant minority of subjects settled on the Pareto inferior distribution (BD,AC), our observations do not confirm C&R's parameter specifications in their model of social preferences.[21] The reason for so many subjects to reject the Pareto improvement (AD,BC) seems to be the considerable payoff difference of 29 Talers, confirming the presence of inequality aversion as modeled by F&S and B&O.

[21] C&R's regression estimates strongly support social welfare preferences ($\rho>\sigma\approx0$) and provide little evidence for difference aversion preferences ($\sigma<0$), cf. pages 838-841, i.e. according to C&R the payoff-unequal Pareto improvement should be chosen over the Pareto inferior more equal allocation. Instead, a significant share of subjects prefers (45,45) over (46,75) which can be accommodated in the C&R framework only by difference aversion preferences ($\sigma<0<\rho<1$).

Observe that either the choice of (BD,AC) and the choice of (AD,BC) is consistent with the CPIES procedure. In R3 of EXP II these two distributions were the only observed outcomes, whereas in R3 in EXP I two subject pairs agreed on the Pareto optimal allocation (AB,CD), a choice which is *not* consistent with the CPIES prediction since (AB,CD) is not a Pareto improvement relative to the equal split distribution (BD,AC), see Table 3. The fact that the CPIES-compatible distribution (AD,BC) was chosen by 19 out of 43 matched pairs whereas the Pareto optimal allocation (AB,CD) was chosen by only 2 of 43 pairs in R3 of EXP I and EXP II combined provides strong evidence for our conclusion that the equal distribution (BD,AC) indeed serves as a focal reference point (a one-tailed binomial test yields a significance level of 0.0001).[22]

Table 8 summarizes the results of treatment *ORD*. As before, the first row gives the percentage of Pareto optimal agreements. The second row quantifies the percentage of agreements on the distribution displaying the minimal difference in the respective ranks in the two players' rankings of bundles. Just as the distribution with minimal payoff difference in the cardinal treatment, this is the distribution that serves as

	EXP III				EXP IV					Total
	R1	R2	R3	R4	R1	R2	R3	R4	R5	
Pareto optimal	86%	92%	100%	91%	78%	88%	100%	96%	100%	92%
Rank Diff Min ("Reference Point")	0%	8%	83%	0%	4%	4%	0%	4%	0%	11%
Rank Diff Min *among* Pareto Opt.	77%	71%	83%	57%	61%	88%	100%	92%	92%	80%
CPIES	91%	79%	83%	91%	100%	92%	100%	100%	100%	93%
Number of Agreements	22	24	23	23	23	24	22	24	24	209

Table 8: Results in treatment ORD (EXP III & IV)

The table quantifies the frequency of final allocations in the different bargaining problems compatible with the criteria given in the rows.

[22] Part of this strong effect might be due to the fact that the sum of payoffs is somewhat larger in (AD,BC) than in (AB,CD).

the reference point for the CPIES procedure in the ordinal treatment. For instance, in R1 of EXP III it is the distribution (AD,BC) since the bundle AD has the same rank in player 1's ranking as the bundle BC in player 2's ranking. The data from R5 of EXP III have been left out in Table 8, since there was no unambiguous reference point, as already noted above. In all other rounds there was a unique reference distribution.[23] The third row gives the percentage of chosen allocations that are Pareto optimal and achieve the minimal rank difference *among all Pareto optimal distributions*.[24] This criterion almost always coincides with the "Rawlsian" criterion of maximizing the welfare of the worst off individual, i.e. in our case of minimizing the maximal rank of a bundle in the player's preference ordering.[25] Observe also that any distribution that minimizes the rank difference among all Pareto optimal distributions either coincides with the reference distribution (in R3 of EXP III) or is a Pareto improvement relative to the reference distribution (in all other examples of EXP III and IV). The fourth row gives the percentage of agreements that are consistent with the CPIES prediction. In EXP III, fewer distributions in which each player received two goods were consistent with the CPIES prediction as compared to EXP IV. As is apparent from the numbers in Tables 7 and 8, the allocations according to the CPIES rule were not chosen randomly. We tested this using a binomial test and found the results to be highly significant; the *p*-values are given in Appendix 2. The CPIES rule can also not be replaced by any of the other criteria – no combination of other criteria explains more than 40% of the variation in the allocation choices based on CPIES in either the ORD or CARD treatment – see Appendix 2 for details.

[23] In R3 of EXP III, the reference distribution was Pareto optimal and chosen in 83% of all cases (cf. Table 8). In all other examples of the ordinal treatment, the reference distribution was not Pareto optimal and chosen only in 3% (=6/186) of all cases.

[24] In R2, R3, R4 of EXP III and in R1, R2, R4 of EXP IV this criterion yields a unique prediction. In the other rounds, there were two rank difference minimizing distributions.

[25] The only exceptions are R4 in EXP III and R3 in EXP IV. In R4 in EXP III there are two "Rawlsian" solutions, namely (BC,AD) and (ABC,D). While both are Pareto optimal, the first of these distributions involves a rank difference of 2, whereas the second distribution represents the lexicographic refinement of the Rawlsian solution with a rank difference of 3. In accordance with the CPIES rule, the first distribution was chosen much more often. In R3 of EXP IV on the other hand, there are two rank-difference-minimizing Pareto optimal distributions, (AC,BD) and (BC,AD); the latter is the unique Rawlsian solution and was chosen much more frequently.

A comparison of the first rows in Tables 7 and 8, respectively, suggests that the failure of Pareto optimality is less prominent in the ordinal than in the cardinal treatment. This is particularly transparent in the comparison of R3 in EXP I (treatment *CARD*) with R2 EXP III (treatment *ORD*). In both cases, there is an unambiguous "equal split" reference distribution, (BD,AC) in R3 of EXP I and (AB,CD) in R2 of EXP III, and exactly one distribution that Pareto dominates this reference distribution; in both cases it is the distribution (AD,BC). In R3 of EXP I, the reference distribution was chosen in 61% (= 14/23) of all cases, whereas in R2 of EXP III only in 8% (= 2/24) of all cases. We conducted a Fisher test and found the differences in behavior in the two treatments to be highly significant.[26]

To summarize, our results confirm the relevance of the CPIES rule as an outcome-oriented criterion in both the cardinal and the ordinal treatment. Our analysis also shows that efficiency in the sense of Pareto optimality plays a more prominent role in the ordinal setting.

Treatment CARD

	Allocation Choice	Payoff Difference	Payoff Sum
Payoff Difference	–0.5354	1.0000	
Payoff Sum	0.2704	0.1475	1.0000
Pareto Optimality	0.1598	0.3145	0.6180

Treatment ORD

	Allocation Choice	Rank Difference	Rank Sum
Rank Difference	–0.2388	1.0000	
Rank Sum	–0.4430	–0.1966	1.0000
Pareto Optimality	0.4021	0.1154	–0.8187

Table 9: Pairwise Correlation Coefficients

The tables are based on the 6 allocations that give 2-good bundles to both individuals.[27] Calculations are based on 1098 observations in the cardinal and 1104 observations in the ordinal treatment. All correlation coefficients are significant at the 1% level.

[26] Specifically, we asked whether subjects chose the Pareto dominated equal split distribution more often in the cardinal treatment than in the ordinal treatment. We tested the equal split distribution both against *all* Pareto optimal choices (case i), and against only the Pareto improvements from equal split (case ii). Moreover, we tested R2 of EXP III both against the data of R3 in EXP I ($p \leq 0.001$ in case i, $p \leq 0.001$ in case ii) and against the aggregated data from the two identical problems in R3 of EXP I and R3 of EXP II ($p \leq 0.001$ in case i, $p \leq 0.005$ in case ii). In each case, the difference in behavior in the ordinal and cardinal treatments is thus highly significant.

[27] Results based on all allocations are qualitatively the same although somewhat noisier; they include the most extreme allocations where one person gets everything or almost everything, and the other person gets little or nothing – these allocations were hardly ever chosen.

This finding is confirmed when considering pairwise correlation coefficients between the choice of an allocation (value: 0 or 1) and Pareto optimality (value: 0 or 1), see Table 9. In the ordinal treatment chosen allocations are more frequently Pareto optimal than in the cardinal treatment. The table also shows that for the chosen allocation the payoff and rank difference are smaller than for the allocations that were not chosen. Moreover, the payoff sum is larger and respectively the rank sum is smaller among the chosen allocations.[28] As in Tables 7 and 8, payoff differences are more closely related to an allocation choice than rank differences. This suggests that when cardinal information is available, comparisons between individuals matter more, whereas when information is only of an ordinal nature, Pareto optimality is more relevant.

Regressions yield similar results (see Appendix 3) and additionally provide direct evidence for the role of CPIES through the inclusion of interaction terms – Pareto optimality combined with payoff or rank difference respectively decreases the likelihood of an allocation being chosen. Therefore, whilst allocations are more likely to be chosen when Pareto optimal, this effect is mitigated if the payoff or rank difference is large, as predicted by CPIES.

4.2 Procedural Aspects

We now take a closer look at the procedural aspects of bargaining behavior, in particular in relation to the CPIES rule. Evidently, not all bargaining problems have equal discriminatory power with respect to the CPIES rule. For instance, in some problems there is no Pareto improvement relative to the reference allocation (EXP I, R4 & R5; EXP II, R1, R2, R4 & R5; EXP III, R3). In all these problems, a clear majority of bargaining partners settled on the reference allocation. While this is clearly in line with the CPIES prediction, it can be counted as evidential support only for the “reference-

In R1 and R4 of EXP III (treatment ORD), the allocation rankings are not monotone w.r.t. bundle size and they were therefore excluded.

[28] Observe that *low* ranks correspond to preferred bundles, so that a negative sign indicates that better ranked allocations were chosen more frequently.

dependent" part of the CPIES rule.[29] To test for the specific "Pareto improvement" part, it is useful to concentrate on specific bargaining problems. Specifically, we will consider R3 of EXP I and EXP II, respectively, in the cardinal treatment, and R2 of EXP III and EXP IV, respectively, in the ordinal treatment. We turn to the cardinal treatment first.

Figure 1 describes how the relative share of the allocations in the agreements evolved over time in EXP I, R3 and EXP II, R3. The bars in the figure describe the relative frequency of allocations in the agreements reached up to the corresponding minute. Thus, e.g., the second bar on the left side means that in EXP I, R3 two thirds of all agreements that were reached in the first three minutes were on the equal distribution (labeled EQ), 17% on the Pareto optimal allocation (AB,CD) which does not represent a Pareto improvement relative to the equal distribution (denoted by PO), and 20% on the allocation (AD,BC) which Pareto improves upon the equal distribution (denoted by PI), and so on. In EXP I, one can see a tendency of the Pareto improving allocation (AD,BC) to gradually gain support: the share of both the agreements on the equal distribution (BD,AC) as well as on the non-improving Pareto optimal allocation (AB,CD) decreases over time. A similar pattern can be discerned in EXP II as well, with the proviso that already in the first minute one pair agreed on the Pareto improving allocation.

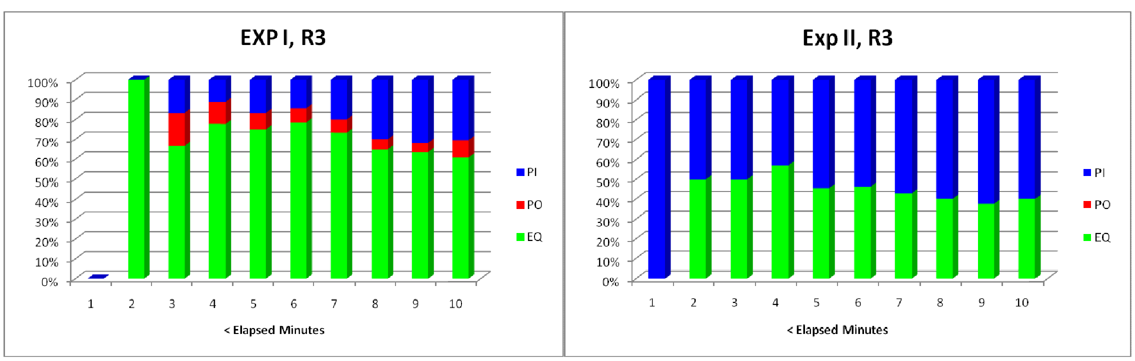


Figure 1: Cumulative final allocations in treatment CARD

The bars describe the relative frequency of allocations in the agreements reached up to the corresponding minute. The Pareto improving allocation predicted by CPIES is denoted by "PI" and the equal reference allocation is labeled "EQ"; the Pareto optimal allocation which does not improve upon the equal distribution is denoted by "PO".

[29] Nevertheless, there is evidence also in the discussions of these bargaining problems that some players *tried* to follow the CPIES rule. They discuss which allocation has the smallest difference in rank and payoffs and mention that they are unsuccessfully looking for improvements.

The temporal pattern revealed by Figure 1 may serve as a first indication of a procedural aspect of the CPIES rule, since evidently the identification of the equal distribution and the possibilities to Pareto improve upon it needs effort and time. However, these numbers are based on final agreements only. To shed light on the principles that guided subjects' bargaining behavior, we analyzed the communication protocols of bargaining partners. The following table quantifies how many pairs of subjects mentioned the equal distribution and the possibility of Pareto improvement, respectively, at some stage in the bargaining process. The tables distinguish between pairs that settled for the equal distribution (EQ) and those who chose the Pareto improvement of the equal distribution (PI).

EXP I, R3

Equal Distribution Mentioned	Final Allocation		Total
	EQ	PI	
Yes	14	5	19
No	0	2	2
Total	14	7	21

Fisher exact test (one-tailed) 0.1000

Pareto Improvement Discussed	Final Allocation		Total
	EQ	PI	
Yes	2	4	6
No	12	3	15
Total	14	7	21

Fisher exact test (one-tailed) 0.0642

EXP II, R3

Equal Distribution Mentioned	Final Allocation		Total
	EQ	PI	
Yes	8	8	16
No	0	4	4
Total	8	12	20

Fisher exact test (one-tailed) 0.1022

Pareto Improvement Discussed	Final Allocation		Total
	EQ	PI	
Yes	2	7	9
No	6	5	11
Total	8	12	20

Fisher exact test (one-tailed) 0.1569

Table 10: Reference to CPIES in the communication protocols (treatment CARD)

The table gives the number of pairs in EXP I, R3 and EXP II, R3 who mentioned the equal distribution and/or discussed the issue of Pareto improvement. Pairs are distinguished according to the final allocation on which they settled (EQ or PI).

Table 10 provides clear evidence that *some* bargaining partners indeed used the CPIES rule (or something close to it) as a guide for their reasoning.[30] For instance, consider the pairs that settled on the equal distribution (14 pairs in EXP I and 8 pairs in EXP II). The entries in the two bottom matrices of Table 10 reveal that four of them, two in EXP I and two EXP II, were aware of the possibility of a Pareto improvement in the bargaining process but rejected it in favor of equality. The entries in the two top matrices are perhaps even more significant. Consider the pairs who settled on the Pareto improving allocation according to CPIES (7 in EXP I and 12 in EXP II). Of these, 5 pairs in EXP I and 8 pairs in EXP II mentioned the equal distribution during the bargaining process; moreover, 6 of these pairs explicitly mentioned the fact that they were both better off than with an equal split in the chat protocols. From this we conclude that the CPIES rule describes the behavior of at least a significant subgroup of our subjects quite accurately in these bargaining problems.

We turn to the ordinal treatment. Some of the scenarios involved several different possibilities to improve upon the reference allocation, i.e. the allocation that minimizes the rank difference. For instance all three allocations chosen in EXP IV, R5 constitute Pareto improvements relative to the reference allocation. In such examples, the focus is not so much on whether the CPIES rule applies or not, but on *which* improvement should be implemented. We therefore concentrate in the following on the ordinal problems EXP III, R2 and EXP IV, R2 which are structurally the most similar to EXP I, R3 and EXP II, R3. The following table summarizes the results.

[30] Table 10 does not consider the two pairs who settled on the non-improving Pareto optimal allocation (AB,CD) in EXP I.

EXP III, R2

Equal Distribution Mentioned	Final Allocation			Total
	EQ	PI	other	
Yes	2	5	1	8
No	0	12	4	16
Total	2	17	5	24

Chi-test 0.0424

Pareto Improvement Discussed	Final Allocation			Total
	EQ	PI	other	
Yes	2	5	1	5
No	0	12	4	19
Total	2	17	5	24

Chi-test 0.2570

EXP IV, R2

Equal Allocation Mentioned	Final Allocation			Total
	EQ	PI	other	
Yes	1	10	0	11
No	0	11	2	13
Total	1	21	2	24

Chi-test 0.2716

Pareto Improvement Discussed	Final Allocation			Total
	EQ	PI	other	
Yes	0	3	0	3
No	1	18	2	21
Total	1	21	2	24

Chi-test 0.6514

Table 11: Reference to CPIES in the communication protocols (treatment ORD)

The table gives the number of pairs in EXP III, R2 and EXP IV, R2 who mentioned the equal distribution and/or discussed the issue of Pareto improvement. Pairs are distinguished according to the final allocation on which they settled (EQ, PI, or other).

Table 11 suggests that the CPIES rule plays a significant role in the ordinal treatment as well. For instance, of the 17 pairs who settled on the Pareto improvement of the equal split (PI) in EXP III five pairs mentioned the equal distribution; similarly, of the 21 pairs who settled on PI in EXP IV 10 mentioned the equal distribution, again a clear indication that equal split serves as reference point even in cases in which it is not chosen.[31] The main difference of the results in the ordinal treatment (EXP III, R2 and EXP IV, R2) on the one hand and of the cardinal treatment (EXP I, R3 and EXP II, R3) on the other hand is that in the former much fewer pairs settled for the equal distribution. We suspect, based on participant comments in the chat protocols, that the rank of a bundle in the preference ordering is used as a proxy for its unknown monetary value and therefore rank

[31] It is perhaps also remarkable that only one of those pairs who did not choose according to CPIES (five pairs in EXP III and two pairs in EXP IV) mentioned the equal distribution, an indication that these pairs either were not aware of the possibility to implement allocations with equal ranks, or that they did not consider this as an attractive solution.

differences are only an imperfect measure for (in)equality. The allocation minimizing the rank difference is no longer the only possible reference point, which in turn favors Pareto optimal outcomes despite possible rank differences.

5. Conclusion

Our experimental results suggest a particular qualitative description of how agents reach agreements in bargaining problems with indivisibilities, the CPIES procedure. The key element is the role of "equal split" as the reference point for the bargaining procedure. Pareto improvements are implemented *provided* that they do not create too much inequality. Indeed, our most striking finding is that a majority of 51% of bargaining partners reject the payoff distribution (46,75) in favor of the Pareto inferior equal split distribution (45,45) (aggregated data from EXP I, R3 and EXP II, R3). This is in contrast to the results of C&R, who "find a strong degree of respect for social efficiency" (p.849). It also conflicts with Kritikos and Bolle's (2001) experiments in which the majority of participants were efficiency- rather than equity-oriented. However, the experiments in these studies consisted of simple dictator games and not of dynamic bargaining games as in our present study. Indeed, inspecting the communication protocols we found strong evidence that procedural aspects play a decisive role in our experiments. For instance, in R3 of EXP I five of the seven pairs who settled on the Pareto optimal distribution (46,75) considered the equal payoff distribution (45,45) at an earlier stage, confirming the explicit reference status of the latter.

In this example, it also seems to matter *who* is the first to propose the Pareto improvement. A detailed analysis of this issue of *path-dependence* is beyond the scope of the present paper, but generally partners seem to agree more easily on the payoff distribution (46,75) if the first individual suggests it to the second individual than *vice versa*. It seems to matter whether the individual suggesting the Pareto improvement benefits more or less than the other person – an issue, we plan to investigate further.

The failure of Pareto optimality due to equity concerns is much less pronounced in the ordinal treatment, even though we do find evidence that the ranks of bundles in the preference orderings serve as substitutes for the unknown monetary payoff. In the ordinal treatment, in some cases there is no longer one clear equal split reference distribution because the inequality associated with any rank difference is unknown. Despite this difficulty, most participants try to establish an "equal split" reference allocation during the bargaining process confirming the key role this reference allocation plays.

References

Bolton, G., Brandts, J., Ockenfels, A. (2005), "Fair Procedures: Evidence from Games Involving Lotteries", *Economic Journal,* 115, 1054-1076.

Bolton, G., Ockenfels, A. (2000), "A Theory of Equity, Reciprocity and Competition", *American Economic Review,* 100, 166-193.

Bolton, G., Ockenfels, A. (2005), "A Stress Test of Fairness Measures in Models of Social Utility", *Economic Theory,* 25, 957-982.

Bolton, G., Ockenfels, A. (2006), "Measuring Efficiency and Equity Motives. A Comment on 'Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments'," *American Economic Review,* 96, 1906-1911.

Camerer, C. (2003), *Behavioral Game Theory – Experiments in Strategic Interaction,* Princeton, Princeton University Press.

Charness, G., Rabin, M. (2002), "Understanding Social Preferences with Simple Tests", *Quarterly Journal of Economics*, 817-869

Engelmann, D., Strobel, M. (2004), "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments", *American Economic Review,* 94, 857-869.

Engelmann, D., Strobel, M. (2006), "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments – Reply to Bolton and Ockenfels", *American Economic Review,* 96, 1918-1923.

Fehr, E., Schmidt, K. (1999), "A Theory of Fairness, Competition, and Cooperation", *Quarterly Journal of Economics*, 114, 817-868.

Fehr, E., Naef, M., Schmidt, K. (2006), "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment", *American Economic Review*, 96, 1912-1917.

Herreiner, D.K., Puppe, C. (2004), “Envy Freeness in Experimental Fair Division Problems”, *Theory and Decision*, forthcoming.

Kalisch, G., Milnor, J., Nash, J., Nering, E. (1953), „Some Experimental n-Person Games“, RAND Corporation Research Memorandum 948.

Kagel, J., Roth, A. (1995), *Handbook of Experimental Economics. Chapter 4*, Princeton, Princeton University Press.

Kritikos, A., Bolle, F. (2001), “Distributional Concerns: Equity- or Efficiency-Oriented?”, *Economics Letters*, 73, 333-338.

Lind, E.A., Tyler, T.R. (1988), *The Social Psychology of Procedural Justice*, New York, Plenum Press.

Roth, A., Malouf, W. (1979), “Game-Theoretic Models and the Role of Information in Bargaining”, *Psychological Review*, 86, 574-594.

Roth, A., Murnighan, K. (1982), “The Role of Information in Bargaining: An Experimental Study”, *Econometrica*, 50, 1123-1142.

Shor, M. (2007), “Rethinking the Fairness Hypothesis: Procedural Justice in Simple Bargaining Games”, Mimeograph.

Tyler, T.R., Lind, E.A. (2000), “Procedural Justice”, in: Sanders, J., Hamilton, V.L. (eds.), *Handbook of Justice Research in Law*, 63-91, New York, Kluwer.

Appendix 1: Matching of Bargaining Partners

The following table shows how individuals were matched in each round of each session. The numbers in the table refer to the 8 individuals. The columns indicate the rounds. Each row shows the bargaining partners in the respective round with the individual in the role of player 1 named first.

R1	R2	R3	R4	R5
1 5	1 8	1 2	7 1	6 1
2 6	2 5	3 4	8 2	7 2
3 7	3 6	5 6	5 3	8 3
4 8	4 7	7 8	6 4	5 4

Table A1: Matching of Bargaining Partners

Each experiment consists of 6 sessions with 8 participants each, i.e. there were 48 participants in each experiment. Each experiment yielded 6 independent observations of up to 4 (and never fewer than 3) choices by the matched pairs.

Appendix 2: Binomial Test of the CPIES prediction

The following table indicates *p*-values for a one-tailed test. The rows indicate the different treatments. The columns show how many allocations ("#") are compatible with the CPIES prediction and how frequently any of them were chosen ("Occ"). The probabilities ("Prob") have been calculated under the assumption that all allocations in which each player receives two goods occur with a probability of 1/6 (ignoring any other allocations that may have been chosen[32]). In R1 and R4 of EXP III the probability assigned to each allocation has to be adjusted to 1/7 because the allocations with two-good bundles span ranks 4-12 (instead of 6-11) and for 7 allocations[33] both bundles are within that range. The same calculation can be done based on all 16 allocations, assuming that they all are equally likely. Those probability values are obviously even lower than the ones shown in the table.

			#	Occ	Prob
CARD	EXP I	R3	2	21/23	$1.1249 \cdot 10^{-8}$
		R4	1	21/22	$8.4333 \cdot 10^{-16}$
		R5	1	21/22	$8.4333 \cdot 10^{-16}$
	EXP II	R1	1	19/24	$2.9488 \cdot 10^{-11}$
		R2	1	22/23	$1.4689 \cdot 10^{-16}$
		R3	2	20/20	$2.8680 \cdot 10^{-10}$
		R4	1	23/24	$2.5536 \cdot 10^{-17}$
		R5	1	22/24	$1.4817 \cdot 10^{-15}$
ORD	EXP III	R1	3	20/22	$3.5396 \cdot 10^{-6}$
		R2	2	19/23	$1.6667 \cdot 10^{-6}$
		R3	1	19/23	$7.2964 \cdot 10^{-12}$
		R4	3	21/23	$1.6561 \cdot 10^{-6}$
	EXP IV	R1	4	23/23	$8.9105 \cdot 10^{-5}$
		R2	2	22/24	$4.0824 \cdot 10^{-9}$
		R3	2	22/22	$3.1866 \cdot 10^{-11}$
		R4	3	24/24	$5.9605 \cdot 10^{-8}$
		R5	3	24/24	$5.9605 \cdot 10^{-8}$

Table A2: p-values for the CPIES prediction

[32] In R5 of EXP I a total of 23 choices were made and in R2 of EXP III a total of 24 choices – however, in each of these two cases, one allocation is outside the range of two-bundle allocations. Taking the additional allocation into account changes the probability for R5 in EXP I to $8.1559 \cdot 10^{-15}$ and for R2 in EXP III to $5.4792 \cdot 10^{-6}$.

[33] The relevant allocations are (AB,CD), (AC,BD), (AD,BC), (BC,AD), (BD,AC), (CD,AB), and (ABC,D).

Appendix 3: Regression Results

The regression table below models the probability of an allocation choice (1 if chosen) as a function of the properties of that allocation. In line with related work, such as C&R, the criteria considered are the payoff difference (PD), payoff sum (PS), and Pareto optimality (PO) in the cardinal allocation problems, and the rank difference (RD), rank sum (RS), and PO in the ordinal allocation problems. The analysis focuses on the six allocations where each individual receives 2 goods.[34] The table reports results of OLS regressions and nested random effects models (GLLAMM[35]); the latter take into account that choices for each matched pair are related (same allocation problem), that choices within the same groups may be related due to the stranger matching scheme across different rounds (Appendix 1), and lastly, they consider the fact that choices within an experiment may be related because all groups faced the same sequence of the same problems.

	Allocation Choice Treatment CARD	
	OLS	GLLAMM
PD	–0.0137 (0.0011)	–0.0110 (0.0012)
PS	0.0040 (0.0003)	0.0091 (0.0010)
PO	0.3851 (0.332)	0.3760 (0.0327)
PD * PO	–0.0097 (0.0014)	–0.1280 (0.0015)
N	1098	1098
F(4,1094)	331.79	
R-squared	0.5482	
Log-Likelihood		-123.3615

	Allocation Choice Treatment ORD	
	OLS	GLAMM
RD	–0.0609 (0.0077)	–0.0344 (0.0101)
RS	0.0080 (0.0012)	–0.0466 (0.0092)
PO	0.6976 (0.0700)	0.5323 (0.0619)
RD * PO	–0.3580 (0.0722)	–0.1447 (0.0159)
N	1104	1104
F(4,1100)	160.88	
R-squared	0.3688	
Log-Likelihood		-236.5628

Table A3: Allocation Choice as a Function of Allocation Parameters

Standard errors are in brackets. All coefficients are significant at the 1% level. In the OLS regressions parameters are jointly significant at the 1% level.[36]

[34]This eliminates two observed allocation choices in both the cardinal and ordinal treatment (EXP I R5, EXP II R1, EXP III R2, EXP III R5) and increases explanatory power. In the ordinal treatment two rounds (EXP III, R1 and R4) were dropped too, because the nature of the rankings in terms of the cardinality of the bundles is quite different there from the other allocation problems.

[35] See http://www.gllamm.org/. Here, a Gaussian function was used and the random effects were nested from matched pairs over groups to experiments. The number of conditions in the cardinal case is 264.1111 (other specifications with similar log likelihoods have much higher numbers of conditions); the only non-negligible random effect variance is the experiment with 0.4202. The number of conditions in the ordinal case is138.1275 (and is the only specification that converges); and again the only non-negligible random effect variance is the experiment with 0.9436.

[36] OLS regressions with robust errors or clustered on the groups yield F and t values that do not change joint or separate significance levels perceptibly.

The main explanatory variable here is Pareto optimality (dummy variable), which, as noted before, has a much stronger impact in the ordinal case. An interesting feature emerges from the interaction effect. The Pareto optimality effect for an allocation choice is mitigated strongly by the payoff or rank difference, i.e. although Pareto optimality strongly and significantly increases the likelihood of choosing an allocation, this is much less the case if there is a large payoff or rank difference – confirming the CPIES analysis. The results here should be interpreted with great care. As is well known, a linear probability model cannot prevent values outside the [0,1] range – we observe such values here, although very close to the interval boundary. A multinomial model would address this; however, it cannot be applied here, since allocations are not comparable across treatments except for the equal-split allocation.

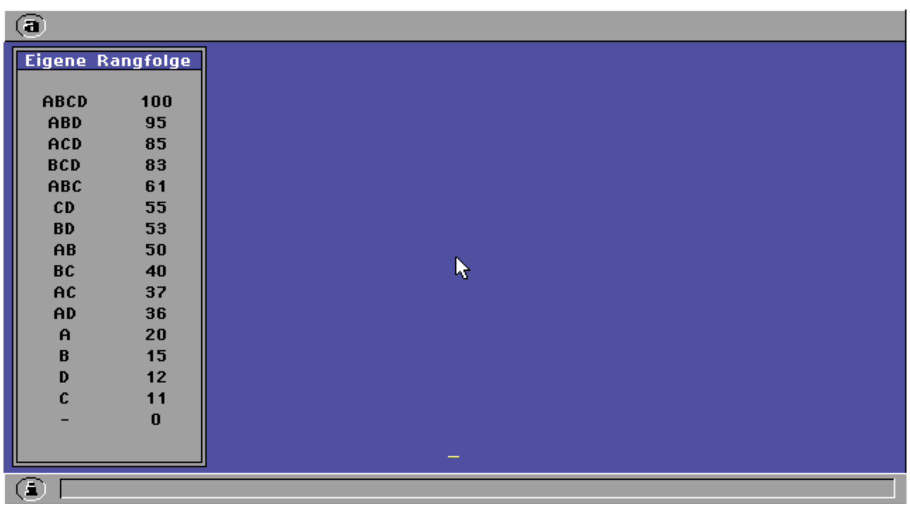
Appendix 4: Instructions and Screen Examples

(The following is a translation of the German instructions for EXP 1 – as close as possible to the German original. The original instructions are available upon request from the authors.)

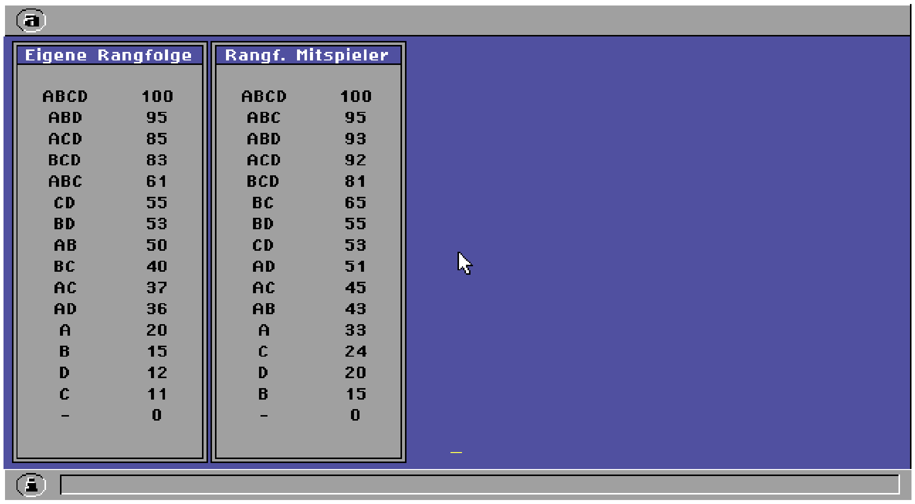
In this experiment you will repeatedly have to distribute several goods between yourself and a partner. The experiment has five independent rounds, each of which you will play with a different partner. In each round you will be given four goods, and you will have to agree with your partner on a distribution of these goods.

There will be four new goods in each round. The goods are referred to as A, B, C and D, respectively. You can think of any kind of object and any kind of division problem. The goods themselves are indivisible, i.e. each good can either be given to you or to your partner. All goods have a positive value. The more goods you receive, the better. However, the value of the goods is different for you and your partner. In each round, we give you a ranking of the bundles of goods in which the values of the bundles are listed in descending order. In each round, you will be given a new ranking. The ranking gives the value of each bundle of goods in Taler (T), our experimental currency. If you agree with your partner on a distribution of goods, you will receive the Taler amount corresponding to your bundle of goods. At the end of the experiment, these Taler amounts will be converted in Deutsche Mark (DM) and paid out to you.

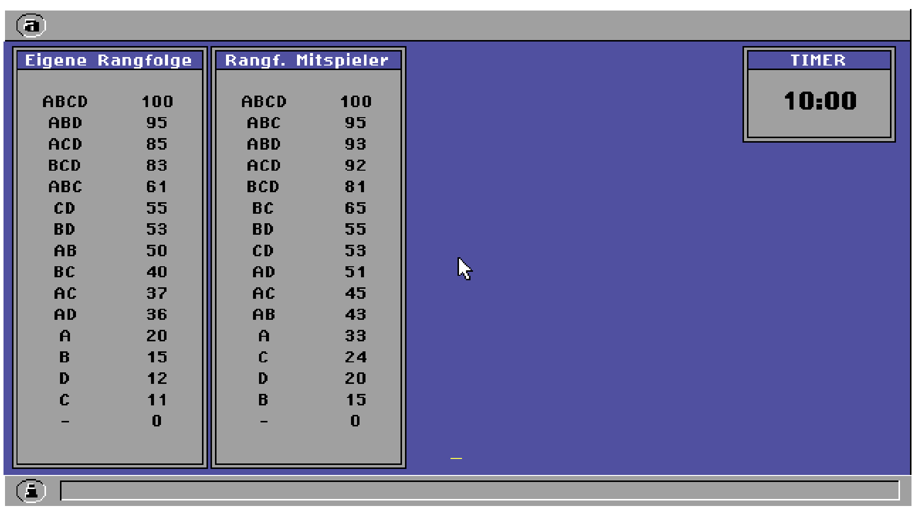
For example, your ranking could look as follows:



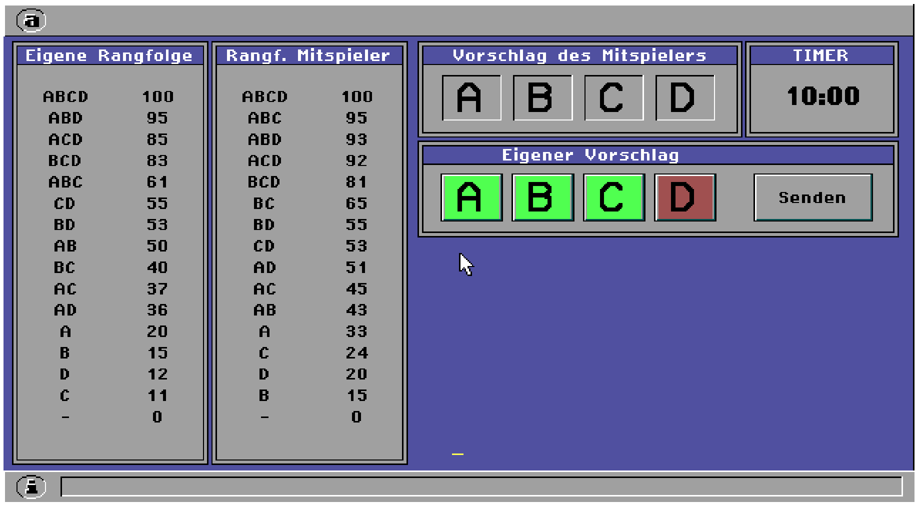
In this case, your most preferred bundle consists of goods A, B, C and D; it is worth T 100. Thus, if you and your partner agreed that he gets nothing and you get all four goods, then you would receive T 100. Your second best bundle is ABD, for which you would receive T 95 if you agreed with your partner that you get ABD and he gets C. Observe that the value of bundles of goods cannot be derived from the values of the single goods. For instance, good C alone is worth T 11 and good D alone is worth T 12, but both goods combined (CD) are worth T 55 to you. It is also possible that a good is worth little when added to another bundle, e.g. the bundle ABD is worth T 95 to you and adding C increases the value of the bundle only to 100 (ABCD), although good C alone is worth 11. In this case, good C does not add much value to the bundle ABD. The goods complement each other in different ways depending on the specific goods with which they are combined. Therefore, for all evaluations in this experiment you have to look at all bundles of goods and not only at the values of single goods. Your partner also gets a ranking of his valuations. On the screen, you will see your partner's ranking next to your own. This may look as follows:



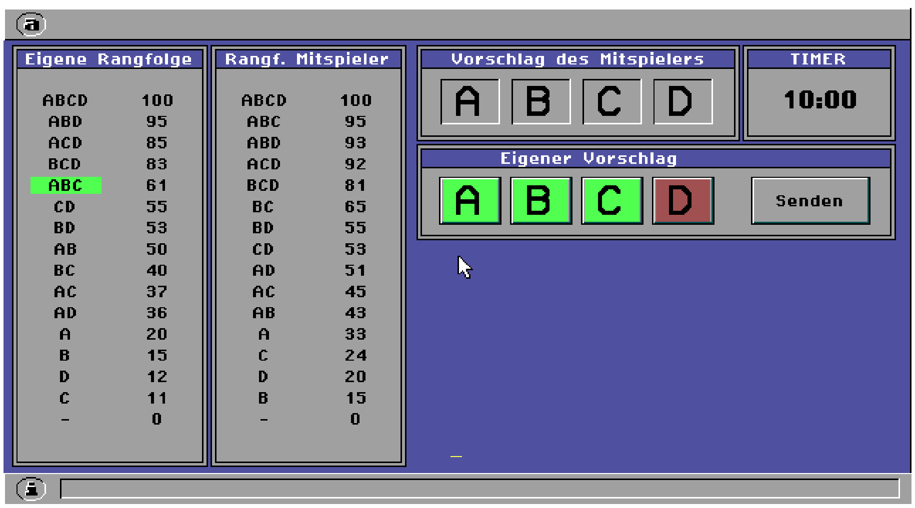
Please start each round by carefully looking at both rankings. The rankings will be different in each round. Each round of this experiment lasts 10 minutes at most. This time is indicated at the top right side of the screen and will be counted down to 0:00 during the round. Within this time span you have to reach an agreement with your partner on who gets which good. If you do not agree within 10 minutes, neither of you will receive anything in this round.



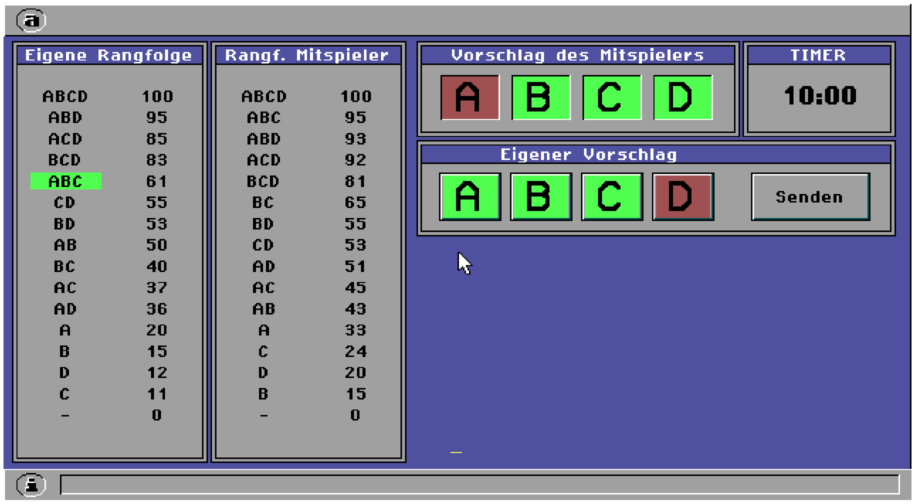
You reach an agreement with your partner by sending him a proposal or by waiting for his proposal. Each of you can make a proposal at the same time. Your partner's proposal appears in the top middle section and your own proposal appears directly beneath. In *both* proposal lines, the goods you get appear in *green*, those received by your partner in *red*. To make a proposal, select the goods you want to receive by clicking on the corresponding buttons, and then send the proposal by clicking on the "send" button.



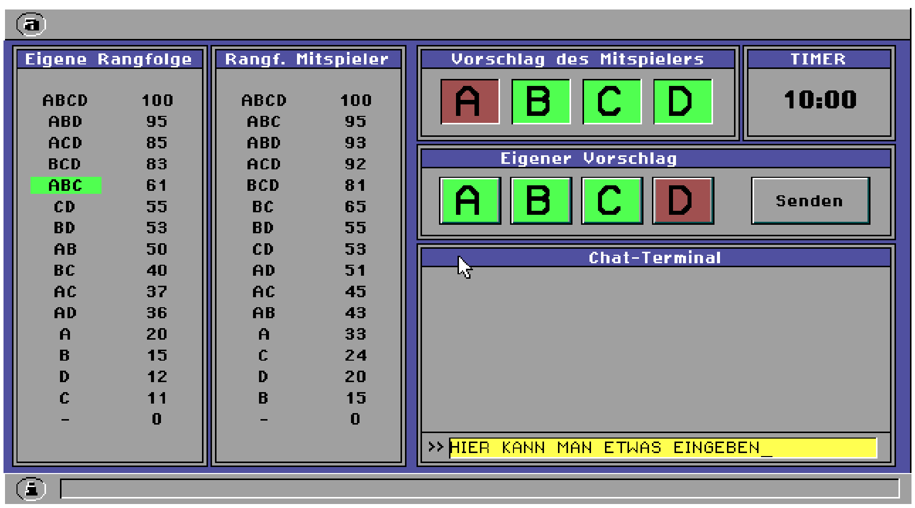
You can change your proposal at any time by clicking on the A, B, C, D buttons. Every click changes the color of the button and therefore moves the good from you (green) to your partner (red) or vice versa. Unless you send your proposal, your partner cannot see your current selection. The most recent proposal you sent can be seen in your ranking on the left – your corresponding bundle is shown in a green box.



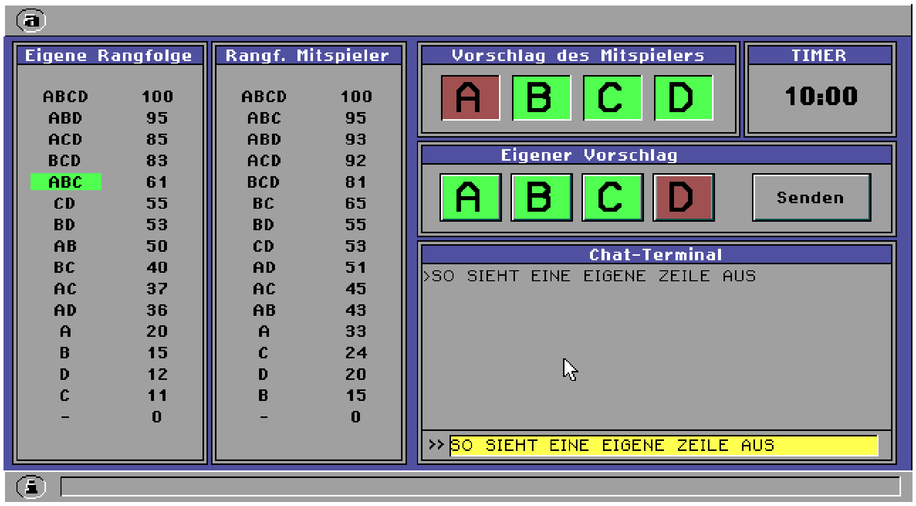
Do not delay sending your proposal because your partner will otherwise not know what you propose. You can change your mind at any time and send a new proposal.



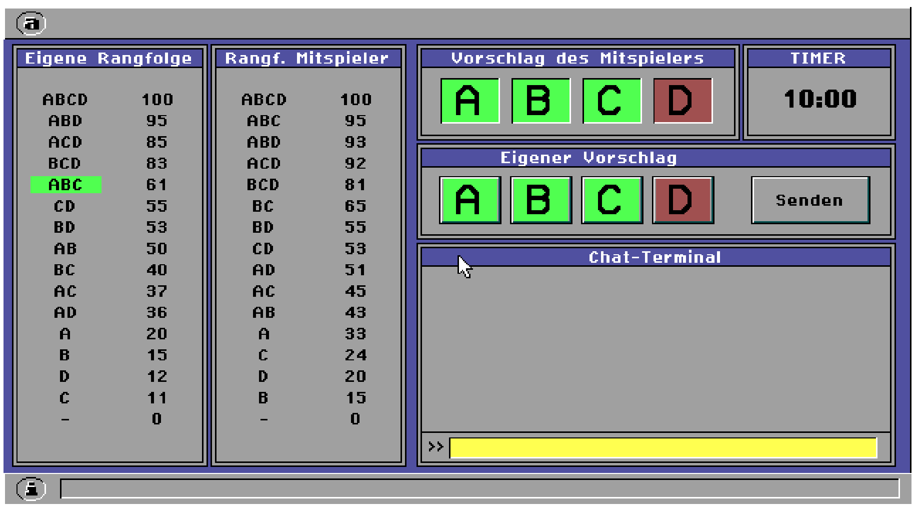
In order to convince a partner to accept your proposal, you can exchange messages in a "chat" window at the bottom right by commenting on your or your partner's proposal. To write in the chat line (max. 80 characters), you have to click on it with the mouse. Press the "enter" key to send a comment. If you want to leave the chat line without writing anything or without sending a comment, you have to press the "Esc" button. If you want to change your proposal after having sent a comment, you will need to leave the chat line first.



Your own comments appear in the chat terminal window with a leading ">" sign; your partner's comments are shown without any additional sign at the beginning.



If the colors of all buttons in your proposal coincide with the colors of the buttons in your partner's proposal, then you have made identical proposals.



You will then be asked whether you want to accept that proposal.

If you and your partner select "Accept" the proposal is accepted and the round is over. If neither or only one of you accepts the proposal, then the round continues, i.e. you can make new proposals or repeat old proposals, and chat. A round is over either if you have both accepted a proposal or if the time limit is reached. If the round has ended before the time limit, you will have to wait until the round is over for all other players – this will be indicated by an acoustic signal. Then, the next round starts for everybody.

At the end of each round you receive the Taler amount corresponding to your bundle of goods. If you did not reach an agreement with your partner you receive no bundle of goods and therefore no Taler amount. The Taler amounts you received will be added over the rounds and converted into DM at the end of the experiment. T 12 equal DM 1.

You will play with a different player in each round of the experiment, hence you *never* play with someone you have already played with. You and your partner *do not know* with whom you play; you will be matched anonymously. What proposals you make, what comments you send, and what bundle of goods you receive in any given round has *no* impact on your or your partner's ranking of bundles, or on the matching of partners in future rounds.

Please do not mention your name and do not make any comments that could reveal your identity. If you violate this rule you will receive no payment!

All relevant information will appear on the screen. A status line at the bottom of the screen indicates the current state of the experiment. Before starting the experiment, you receive a number that corresponds to your computer terminal and you will be paid at the end based on your number.

Do you have any questions?
Please switch off your cell phones for the duration of the experiment.
Thank you for your cooperation.
Good Luck.