The False Consensus E¤ect Disappears if Representative Information and Monetary Incentives Are Given¤

Dirk Engelmann[y] & Martin Strobel[z]
Humboldt-Universität zu Berlin

July 8, 1999

Abstract

In this paper we present an experiment on the false consensus effect. Unlike previous experiments, we provide monetary incentives for revealing the actual estimation of others' behavior. In each session and round sixteen subjects make a choice between two options simultaneously. Then they estimate the choices of a randomly selected subgroup. For half of the rounds we provide information about other subjects' choices. There we …nd no false consensus e¤ect. At an aggregate level, subjects signi…cantly underweight rather than overweight their choices. The results are di¢cult to interpret for the rounds where we do not provide information.

JEL-Classi…cation: C91, D83, D84

¤We thank Werner Güth, Thomas Heße, Ste¤en Huck, Ulrich Kamecke, Daniel Klapper, Manfred Königstein, Hans Normann and Wolfgang Scholl for helpful comments and Klaus Abbink and Abdolkarim Sadrieh (1995) for providing RatImage which was used to implement the experiment software. Financial support by the Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 373 and by the Deutsche Girozentrale - Deutsche Kommunalbank is gratefully acknowledged.

[y]Institut für ö¤entliche Wirtschaft und Wirtschaftspolitik, Humboldt-Universität zu Berlin, Spandauer Str. 1, D-10178 Berlin, Germany, FAX: +49-30-2093 5787, phone: +49-30-2093 5788, e-mail: engelman@wiwi.hu-berlin.de

[z]Institut für Wirtschaftstheorie III, Humboldt-Universität zu Berlin, Spandauer Str. 1, D-10178 Berlin, Germany, FAX: +49-30-2093 5704, phone: +49-30-2093 5724, e-mail: strobel@wiwi.hu-berlin.de

1 Introduction

The 'false consensus e¤ect' has frequently been reported in social psychological research. The term was …rst used by Ross et al. (1977). Mullen et al. (1985) report 115 studies which show a false consensus e¤ect. It has also been observed in several experiments conducted by economists, although these were not designed to test for this e¤ect.[1] Most of the studies employed a de…nition similar to that used by Mullen et al. (1985):

> False consensus refers to an egocentric bias that occurs when people estimate consensus for their own behaviors. Speci…cally, the false consensus hypothesis holds that people who engage in a given behavior will estimate that behavior to be more common than it is estimated to be by people who engage in alternative behaviors.

In 1989 Dawes (1989 & 1990) realized that this type of de…nition does not justify the label 'false'. He argues that it is perfectly rational to use the information about one's own decision, behavior, or whatever else is being considered in the same way as the information about any other randomly selected sample of size one. The e¤ect is only false if too much weight is assigned to one's own decision. We will therefore refer to the e¤ect as de…ned above as a consensus e¤ect and use the following de…nition of a false consensus e¤ect which is both stricter and more appropriate:

> A (truly) false consensus e¤ect is considered to be present if people, when forming expectations concerning other people's decisions, weight their own decision more heavily than that of a randomly selected person from the same population.

In contrast to the de…nition cited above this implies that a false consensus e¤ect can be considered as a bias in a Bayesian framework. It would improve our ability to predict human behavior greatly if we could discover systematic deviations from rational expectations, like the false consensus e¤ect, and if we could quantify them. The inclusion of a false consensus e¤ect in game theoretic models can yield interesting insights and results (see e.g. Engelmann, 1998). Whether the false consensus e¤ect can be shown to be present even in terms of the stricter de…nition is therefore of interest to economists.

A reevaluation of several studies using the strict and appropriate de…nition has shown that most of the studies did not …nd such a (truly) false

[1] See for example Selten & Ockenfels (1998) or Jacobsen & Sadrieh (1996).

consensus e¤ect, but that they quite often found an e¤ect in the opposite direction (see Dawes, 1990). Sherman et al. (1984) …nd a (truly) false consensus e¤ect only when the self-image is threatened. Krueger and Clement (1994) …nd a false consensus e¤ect even in the presence of other information. So when the strict de…nition of the false consensus e¤ect is applied, empirical results are rather ambiguous.

Furthermore, most studies do not even allow testing for a (truly) false consensus e¤ect since they do not provide any option for comparing how the knowledge about one's own decision is weighted relative to information about other people's decisions. Often the only attempt to evaluate the falsity of the e¤ect is to assume a uniform prior belief about possible decisions and then compare the posterior belief actually expressed to what would be obtained by rational Bayesian updating using a sample of size one (see Dawes, 1989, and Krueger and Clement, 1994, for details). However, we consider the assumption of a uniform prior belief in this context totally arbitrary and unjusti…able.

Hence we set out to conduct an experiment which, without needing to use unjusti…able assumptions or sophisticated statistical methods, allows a simple test to be made to see if knowledge about the own decision is weighted more heavily than information about other people's decisions.

Our experiment also avoids further shortcomings of previous experiments. One of these is that the information, if provided, was often taken from some general survey. The latter is problematic if, for example, students' behavior is what is being estimated and the subject herself is a student. In this case it is justi…able that she puts more weight on her own behavior than on the information provided. To avoid this problem we provided information from the same population whose behavior was to be estimated later.

From the perspective of an experimental economist, most social psychological experiments on the false consensus e¤ect are also unsatisfactory for methodological reasons. This is primarily due to the fact that no incentives for revealing the true beliefs are provided. Our study (like that of O¤erman et al. (1996) who examine the presence of a false consenus e¤ect in a public good experiment) di¤ers by providing clear (monetary) incentives for revealing the true beliefs. The relevance of incentives in this context is demonstrated by O¤erman et al. (1996). In a post-experiment questionnaire 50% of their subjects stated that they would have answered di¤erently if no incentives had been provided.

Furthermore, in many of the social psychological experiments where information about others people's decisions is provided, this information is rigged and the subjects are clearly deceived. These experiments are therefore unsatisfactory. Apart from general concerns about this practice which we do

not want to discuss at this point (see Camerer (1996) for a general discussion), it is quite obvious that, in some studies, subjects might have become suspicious about this information and thus might have discarded it. In the present study subjects were not deceived.

16 rounds were played in each session of the present experiment. In each round subjects …rst had to choose between two options and then estimate how some of the other subjects in the same session had chosen. They were informed before the estimation step about the …rst step choices of the remaining subjects in half of the rounds. Monetary incentives for true revelation of both preferences and beliefs where provided, since payo¤s were based on the number of correct estimates and transformed through the …rst step choices of half of the rounds. (See the following section for details.)

We found a clear consensus e¤ect both when we provided information and when we did not as did most studies of the false consensus e¤ect (see e.g. Mullen et al.,1985). Where we provided information about other subjects' decisions, the information was clearly used. Although subjects de…nitely assigned some positive weight to their own decision when estimating that of other subjects in the same session, on average this weight was even lower than that assigned to the decisions they were informed about. Hence the consensus e¤ect we found was not a false consensus e¤ect. In the rounds where we did not provide representative information testing for a (truly) false consensus e¤ect is problematic, since the e¤ect of the own decision on the estimation cannot be compared with the e¤ect of representative information simply because there isn't any. Consequently, our results for these items are di¢cult to interpret.

In the next section we describe the experimental design. Section three contains the experimental results, followed by the conclusion.

2 Experimental Design

In the present experiment 16 subjects played 16 rounds simultaneously on a computer network.[2] Each round consisted of three steps. In the …rst step each subject chose between two options of an item (Decision 1). One item, for example, was the question: "What do you eat more frequently, oranges or bananas?" Another allowed the subjects to donate DM 1,- to the non-government organization "amnesty international". (See Appendix C for the complete list.) The second step began with the random assignment of four (di¤erent) subjects to each individual subject. In half of the rounds

[2] That the number of subjects equals the number of rounds is a mere coincidence and of no signi…cance to the experiment.

		With Information	Without Information
Payo¤ Relevant	Morally Relevant	Rich-Poor (6) Donation Chapel (15)	Public Good (4) Donation ai (9)
	Morally Irrelevant	Lottery (7) Time Preference (10)	Payo¤ System (3) Stamps (14)
Not Payo¤ Relevant	Morally Relevant	Free-Riding (2) Waiter (16)	Tax Evasion (5) Exam (12)
	Morally Irrelevant	Fruit (1) Months (13)	Writing (8) Elevator (11)

Table 1: Decision items 2 £ 2 £ 2 design. The numbers in brackets correspond to the order in which the items were presented to the subjects. For details see Appendix C.

subjects were then informed about these four subjects' Decisions 1 without identifying them. For the other rounds subjects received no information. The task in the third step was to estimate the remaining eleven subjects' Decisions 1 (this is called Decision 2). Subjects were informed that the random matching was conducted independently for each subject and for each round. This procedure ensured that the information subjects were provided with was from the same population as the subjects whose decisions they were to estimate. No feedback about the success in Decision 2 was provided before the experiment was completed so as to prevent learning across items.

We implemented a 2 £ 2 £ 2 design with respect to the items (see Table 1). Items di¤er in terms of payo¤ relevance, moral relevance, and whether the subjects received information about the decisions of the four other participants.

The distinction between "with information" and "without information" allows us to compare the traditional way of measuring the (false) consensus e¤ect with the approach we consider more appropriate. The traditional way compares estimates given in Decision 2 by subjects who chose one option in Decision 1 to estimates given by subjects who chose the other option in Decision 1 (see e.g. Ross et al., 1977). This does not require information. The more appropriate way is to compare the weights given to the own Decision 1 and to that of a randomly chosen person in the same population when making Decision 2.

The distinction between morally relevant and irrelevant items gives us the opportunity to investigate whether the (false) consensus e¤ect has an ego-defensive function, for example: "I do free-ride on public transport, but everybody does." If this were the case, the e¤ect should be larger for morally

relevant items than for morally irrelevant items. Sherman et al. (1984) …nd evidence for this hypothesis which is doubted by Alicke and Largo (1995).

An alternative explanation for the phenomenon usually labeled false consensus e¤ect could be that it is not a distortion of the expectations about other people's decisions but rather an adaptation towards the perceived decision of the majority. This e¤ect could be based on a desire to comply with perceived social norms. In the present context, this means that subjects adapt to assumed group behavior and norms as expressed in Decision 2 when making Decision 1. If this were the case, the e¤ect should be smaller for the payo¤ relevant items because we provide clear monetary incentives for revealing subjects' true preferences.

Each cell in Table 1 contains two items. The numbers in brackets correspond to the order in which the items were presented to the subjects. The payo¤s to the subjects depended primarily on the number of exact hits in Decision 2 in all rounds. They were also in‡uenced by the payo¤ relevant Decisions 1. Thus we provided clear monetary incentives for the true revelation of both the true estimates of other subjects' decisions and of the own preferences regarding Decision 1.

The protocol of the experiment as well as a translated version of the instructions can be found in appendices A and B.

3 Experimental Results

Altogether we conducted six sessions, each with 16 participants. The data from the …rst session (pilot session) are not included in the analysis, since we slightly changed the setup and some items. Hence, we have data on 80 subjects from …ve independent observations. Participants were mainly economics and business administration students at Humboldt-Universität. Each session took roughly one hour. The average payo¤[3] was about DM 21.12.

3.1 Analysis of Items with Information

The main results for the items with information are summarized in Tables 2, 3 and 4. Table 2 deals with the presence of a consensus e¤ect. Here, options in Decision 1 of each round are labeled with A and B for simplicity. The option that appeared on the left side of the screen is labeled option A, that

[3] Participants could choose to get their payo¤ in stamps. If they did, we counted the cash equivalent of the stamps, which was DM 1.10.

Info Set	4A-0B		3A-1B		2A-2B		1A-3B		0A-4B	
Decision	A	B	A	B	A	B	A	B	A	B
Item	#	#	#	#	#	#	#	#	#	#
	;	;	;	;	;	;	;	;	;	;
	C/N		C/N		C/N		C/N		C/N	
Fruit	0	3	4	6	14	13	6	9	3	22
	-	8.7	7.5	4.8	5.7	4.8	3.8	3.2	3.7	1.7
	-		C		C		C		C	
Free-Riding	2	1	10	6	7	15	9	18	6	6
	10.5	9	6.5	7.2	6.6	4.6	3.9	3.7	2.7	2.7
	C		N		C		C		N	
Rich-Poor	0	0	2	8	6	10	9	20	7	18
	-	-	7.5	5.4	5.8	3.8	4	1.9	2.9	0.4
	-		C		C		C		C	
Lottery	0	0	0	1	0	6	5	28	7	33
	-	-	-	6	-	3.8	3.8	2.0	2.3	0.8
	-		-		-		C		C	
Time Preference	1	0	3	4	8	12	13	23	3	13
	10	-	9	6.5	6.6	4.8	3.6	3.1	3.3	0.6
	-		C		C		C		C	
Months	1	3	5	7	3	14	16	16	5	10
	10	7	8.2	5.7	6	4.7	3.3	2.9	1.4	1.6
	C		C		C		C		N	
Donation Chapel	1	0	1	8	10	18	14	19	4	5
	11	-	8	5.8	5	4.6	3.7	3.0	3.5	0.2
	-		C		C		C		C	
Waiter	2	0	2	4	11	10	11	26	3	11
	10	-	9	5.5	5.5	4.3	3.4	1.9	1.7	0.5
	-		C		C		C		C	

Number of subjects chosing A (left) or B (right) for the item in the corresponding row and receiving the information in the corresponding column

; Average estimate given by these subjects of how many of eleven subjects chose A for the corresponding item

C Di¤erence of estimates directly above corresponds to a consensus e¤ect

N Di¤erence of estimates directly above does not correspond to a consensus e¤ect

Table 2: Test for consensus e¤ect (own decision not considered as information)

on the right side is labeled option B.[4] In the rest of this section we will refer to the subjects who chose option A in Decision 1 as 'A-subjects' and the subjects who chose option B in Decision 1 as 'B-subjects'.

As in most "classical" studies of the false consensus e¤ect, in Table 2 we simply compare the estimations of how many of eleven subjects chose option A given by the A-subjects with those given by the B-subjects. We consider a consensus e¤ect present if the …rst estimate is higher than the second, because this means that at least one of the groups estimates that the consensus for the own decision is greater than is justi…ed by the information given. Since the information provided clearly matters (as will be explained below), it is only reasonable to apply this comparison to each set of subjects who received the same information separately.

The rows in Table 2 refer to the di¤erent items while the columns refer to the di¤erent information sets that the subjects were provided with before Decision 2. Each cell contains …ve entries. These are the number of A-subjects (upper left) and B-subjects (upper right) who received the corresponding information set (aggregated over all …ve sessions), the average estimate of how many of eleven other subjects chose option A given by the A-subjects (middle left) and by the B-subjects (middle right) and an indicator (bottom line) which is C if a consensus e¤ect is present and N otherwise. Consider for example the item "Free-Riding" and the information set "1A-3B". The entries in this cell show that 9 A-subjects and 18 B-subjects received the information "one other subject chose A and three subjects chose B" and that their average estimates were 3.9 and 3.7, respectively. Thus a consensus e¤ect is present which is indicated by the C in the bottom line.

3.1.1 The E¤ect of the Information

Examining Table 2, the …rst important thing we see is a clear e¤ect of the information given. The less A's the subjects have in their information sets, the lower their estimates of the number of subjects who chose option A. To see this, consider any item and either the A-subjects or the B-subjects, for example the item "Waiter" and the A-subjects. Those subjects who had four A's in their information set on average estimated that 10 out of 11 subjects chose option A. The estimates strictly decrease with the number of A's in the information set to 9, 5.5, 3.4, and eventually to 1.7 for the information set containing no A's. This strict decrease in A estimates is violated only twice. For the item "Fruit", the B-subjects on average estimated 4.8 choices of option A both if they had two or three A's in their information set. For the

[4] To subjects, options always were named by an abbreviation of the respective meaning (see Appendix C), not by these labels.

item "Free-Riding", estimates increase from 6.5 by the A-subjects who had three A's in their information set to 6.6 by those who received just two. All other comparisons between two cells are in the direction suggested. Hence the information given to subjects clearly a¤ects their estimates as expected.

The observations within each session are not independent and thus it is not possible to apply straightforward tests to the aggregate data. An alternative method is to treat each session as an independent observation and consider each item separately. Any comparison between two information conditions which shows a higher number of A estimates for the subjects with more A's in their information set than for those with fewer A's (and the same own decision) contradicts the hypothesis that information has no in‡uence. Hence we simply count the results of all these comparisons and consider one session to be a hit for this item if the number of positive di¤erences is larger than the number of negative di¤erences. This occurs under the hypothesis that information has no in‡uence with probability of less than $\frac{1}{2}$ (since there is a positive probability for a draw). Therefore the probability for a hit in all …ve sessions is smaller than $\frac{1}{32}$ = 3:125% and hence in this case the hypothesis can be rejected for the item at the 5%-level. Using this procedure we …nd a signi…cant e¤ect of the information for six of the eight items (data not reported). For the remaining two items ("Fruit" and "Time Preference") there are four hits and the e¤ect of the information is thus in the same direction but not signi…cant. Thus the overall e¤ect of information on the estimates of other subjects' behavior is well established.

3.1.2 The Consensus E¤ect

The second important observation that can be drawn from Table 2 is the presence of a consensus e¤ect. Of 32 cells where a comparison between the estimates given by the A-subjects and the B-subjects could be made, 29 show an e¤ect as predicted by the consensus hypothesis. Of the three remaining cells, there are one draw (item "Free-Riding", information set "0A-4B") and two di¤erences in the opposite direction, one small (item "Months", information set "0A-4B") and one somewhat substantial (item "Free-Riding", information set "3A-1B"). So there is a very clear indication of a consensus e¤ect in the present study.

As noted above, for strict statistical tests we have to treat the sessions as independent observations instead of the decisions. For any item, any session, and any information condition, we compare the estimates given by the A-subjects with those given by the B-subjects. We count a session as a hit for this item if for more information conditions the di¤erence between average estimates given by A-subjects and by B-subjects is positive rather

than negative. Under the hypothesis that the own decision is irrelevant to the estimate, i.e. the absence of a consensus e¤ect, the probability for a hit is less than $\frac{1}{2}$ (since the probability for a draw is positive). Corresponding to the analysis of the e¤ect of information, there is a signi…cant consensus e¤ect at the 5%-level for an item if there is a hit in all …ve sessions. Using this procedure we …nd a signi…cant consensus e¤ect for two items ("Fruits" and "Waiter"). Of the remaining items four have four hits, one three hits and two draws and one two hits, two draws and one miss (data not reported). If the probability for a draw is taken into account there is a signi…cant consensus e¤ect (5%-level) for two further items ("Rich-Poor" and "Lottery"). Thus the evidence for the consensus e¤ect is rather strong, given that several draws and misses are caused by cells with only one subject so that the variance across cells is very high.

Although we are aware of the fact that we cannot legitimately treat the di¤erent items in one session as independent, it is illustrative to ignore this problem. If we considered each item in each session as independent observation and de…ned hits as above, we would obtain 31 out of possible 40 (eight items times …ve sessions) hits. This would allow the hypothesis that there is no consensus e¤ect to be rejected at a 0:5%-level. Since the possible in‡uence of the dependence of the items is not clear, we consider this as at least a clear indication of the presence of a consensus e¤ect although it is not hard evidence by strict standards.

The two conclusions drawn from Table 2 show that the estimates subjects give for the choices of other subjects are clearly in‡uenced by both the information they receive and their own choice in Decision 1.

3.1.3 The False Consensus E¤ect

As was argued in the introduction, to examine the false consensus e¤ect it must be acknowledged that it is rational to use the information concerning the own decision in the same way as the information about the decision of any randomly selected subject from the same population. Hence, in the present study subjects have the information about …ve subjects and not just four. Thus to decide whether there is a false consensus e¤ect we have to compare those subjects who made di¤erent choices in Decision 1 but received the same information including that about their own decision. An A-subject who received the information set "2A-2B" should expect more A choices than a B-subject who received the same information. Therefore, a comparison should not be made between these two subjects. Instead, the …rst should be compared to the B-subjects who received the information set "3A-1B" since both have the total information that three subjects chose A and two

Info Set	4A-1B		3A-2B		2A-3B		1A-4B	
Decision	A	B	A	B	A	B	A	B
Item	#	#	#	#	#	#	#	#
	;	;	;	;	;	;	;	;
	F/N		F/N		F/N		F/N	
Fruit	4	3	14	6	6	13	3	9
	7.5	8.7	5.7	4.8	3.8	4.8	3.7	3.2
	N		F		N		F	
Free-Riding	10	1	7	6	9	15	6	18
	6.5	9	6.6	7.2	3.9	4.6	2.7	3.7
	N		N		N		N	
Rich-Poor	2	0	6	8	9	10	7	20
	7.5	-	5.8	5.4	4	3.8	2.9	1.9
	-		F		F		F	
Lottery	0	0	0	1	5	6	7	28
	-	-	-	6	3.8	3.8	2.3	2.0
	-		-		N		F	
Time Preference	3	0	8	4	13	12	3	23
	9	-	6.6	6.5	3.6	4.8	3.3	3.1
	-		F		N		F	
Months	5	3	3	7	16	14	5	16
	8.2	7	6	5.7	3.3	4.7	1.4	2.9
	F		F		N		N	
Donation Chapel	1	0	10	8	14	18	4	19
	8	-	5	5.8	3.7	4.6	3.5	3.0
	-		N		N		F	
Waiter	2	0	11	4	11	10	3	26
	9	-	5.5	5.5	3.4	4.3	1.7	1.9
	-		N		N		N	

Number of subjects chosing A (left) or B (right) for the item in the corresponding row and receiving the information in the corresponding column (including their own decision)

; Average estimate given by these subjects of how many of eleven subjects chose A for the corresponding item

C Di¤erence of estimates directly above corresponds to a false consensus e¤ect

N Di¤erence of estimates directly above does not correspond to a false consensus e¤ect

Table 3: Test for (truly) false consensus e¤ect (own decision considered as information)

chose B ("3A-2B"). The second should be compared to those A-subjects who received the information set "1A-3B" since both have the total information set "2A-3B". Only if in these comparisons the A-subjects estimate a higher number of A choices than the B-subjects, would there be an indication for the presence of a false consensus e¤ect because this would imply that greater weight was assigned to the own decision than to that of the subjects about whom information was given.

Table 3 is structured similarly to Table 2, it only has to be noted that the own decision is included in the information set. For example, in the column marked "3A-2B" we note for each item the number and average estimate of A-subjects who received the information set "2A-2B" (upper and middle left) and the number and average estimate of B-subjects who received the information set "3A-1B" (upper and middle right). If, among the subjects with the same total information, the A-subjects on average estimated more A choices than the B-subjects, we note an "F" for "false consensus present" in the bottom line of the cell, otherwise an "N" for "no false consensus present".

A brief inspection of Table 3 reveals that the situation is far less clear than that in Table 2. From 26 cells which allow a comparison, only eleven show a di¤erence between average estimates in line with the false consensus hypothesis, whereas 13 show a di¤erence in the opposite direction and two show no di¤erence. Furthermore, among the eleven cells with a di¤erence corresponding to a false consensus, there are six for which the di¤erence is small (i.e. < 0:5) and six which are based on a very small sample (i.e. less than …ve subjects for at least one group), including three cells with both these defects. This leaves just two cells which show a substantial di¤erence between average estimates corresponding to a false consensus and are based on a larger sample ("Fruits, 3A-2B" and "Rich-Poor, 1A-4B").[5] On the other hand, there are ten cells which show a di¤erence larger than 0:5 in the direction opposite to that of a false consensus and are based on at least …ve subjects per group ("Fruit, 2A-3B", "Free-Riding, 3A-2B, 2A-3B, 1A-4B", "Time Preference, 2A-3B", "Months, 2A-3B, 1A-4B", "Donation Chapel, 3A-2B, 2A-3B", "Waiter, 2A-3B").

A separate analysis of the single items reveals that only for the item "Rich-Poor" the results from all information sets taken together are more in favor of a false consensus e¤ect than the opposite. The possible conclusion that a false consensus e¤ect might be more likely in the respective category ("morally relevant" and "payo¤ relevant") is not sustained, since the other

[5] This result does not in principle depend on the speci…cation of a small di¤erence and a small sample. For example, only four cells ful…ll the weaker requirement of a di¤erence of at least 0:4 and a minimal number of 4 A- and B-subjects.

item from the same category ("Donation Chapel") shows di¤erences which are clearly contrary to this conclusion. Also, if this conclusion were true, the items from the category "morally relevant" and "not payo¤ relevant" ("Free-Riding" and "Waiter") should exhibit a false consensus at least as strongly. But they clearly do not, indeed, just the contrary.

Altogether there is stronger indication of an e¤ect opposite to a false consensus e¤ect, i.e. that subjects weight the information about their own decision less than than they do the information they have been given about other subjects' decisions.

This conclusion is con…rmed by an analysis of the single items for each separate session. As with the consensus e¤ect, we count a session as a hit if for more information conditions the di¤erence between average estimates given by A-subjects and B-subjects is positive rather than negative. For none of the items are there more than two (out of possible …ve) hits. But there is also no signi…cant e¤ect in the other direction. Altogether, there are eleven hits, eleven draws and 18 misses. Hence this analysis to some degree supports the conclusion of an e¤ect opposite to a false consensus.

3.1.4 Regression Analysis

The tendency against a false consensus e¤ect is shown more clearly by a regression analysis. Assume the simple linear regression model

$$E = \bar{}O + °I + ±;$$

where E = # of A's estimated, I = # of A's in the information set and O = # of A's in the own decision, i.e. O = 1 if the own decision is A and O = 0 otherwise. If subjects weight their own decision like they do the information about other subjects' decisions, then $\bar{}$ equals °. A false consensus e¤ect corresponds to $\bar{} > °$ and the underweighting of the own decision to $\bar{} < °$: The estimated coe¢cients are shown in table 4.

The …rst column contains the coe¢cients for the regression over all sessions, the …rst row those over all items. Over all sessions and all items, $\bar{}$ is clearly smaller than ° and hence the own decision is weighted less than information about other subjects. The same holds for any single session over all items. If the own decision were weighted more than, or equal to, the information about other subjects' decisions, the probability that $\bar{} < °$ for a single session would be smaller than $\frac{1}{2}$: Since the sessions are independent, the probability that $\bar{} < °$ for all …ve sessions is smaller than $\frac{1}{32} < :05$. Consequently, the hypothesis that subjects overweight their own decision relative to other subjects' decisions (or that they weight it equally) can be rejected at

session	all	2	3	4	5	6
coe¢cients	¯ °	¯ °	¯ °	¯ °	¯ °	¯ °
items						
all	1.33 1.73	1.44 1.72	1.51 1.78	0.85 1.74	1.35 1.47	1.42 1.83
morally relevant	1.39 1.75	0.36 2.06	1.77 1.87	1.29 1.64	2.00 1.39	1.19 1.66
not morally rel.	1.28 1.73	2.32 1.49	1.36 1.85	0.43 1.84	0.54 1.56	1.57 1.95
payo¤ relevant	1.66 1.78	1.15 1.69	1.97 1.56	1.13 2.13	1.62 1.41	2.25 1.90
not payo¤ rel.	1.01 1.67	1.64 1.58	1.00 1.91	0.37 1.28	1.16 1.52	0.56 1.75

Table 4: regression coe¢cients (¯: own decision, °: information), session 1 (pilot) is not reported.

a 5%- level. Hence we obtain a signi…cant e¤ect opposite to a false consensus e¤ect.

The comparison of the morally relevant to the morally irrelevant items allows us to test whether a potential false consensus e¤ect has an ego-defensive function. Over all sessions there is a tendency in the direction corresponding to this hypothesis, namely the relative weight given to the own decision compared to the information is slightly higher for the morally relevant items than for the morally irrelevant items. However, we do not consider this to indicate that this hypothesis is correct because the weight given to the own decision is still smaller than that given to the information. The fact that the e¤ect contrary to a false consensus is smaller cannot be considered an indication that the (non-existing) false consensus e¤ect is larger for these items.

The comparison between payo¤ relevant and not payo¤ relevant items allows us to test the hypothesis that a possible false consensus e¤ect is not based on the in‡uence of the own decision on the estimation of others' decisions, but rather on an adaptation of the own decision towards the estimated majority. If this were the case, then the e¤ect should be smaller for the payo¤ relevant items since the incentives for expressing the true preferences work in the opposite direction. Hence the relative weight of the own decision should be smaller for the payo¤ relevant decisions. We …nd just the opposite. We only have a speculative explanation for this result. These items are further from everyday experience than the not payo¤ relevant items. This could re-

sult in a reasoning process which is based less on general expectations and the given information and more on introspection.

Following our general argument that incentives matter, it may be argued that the analysis should be restricted to the payo¤ relevant items. Then, over all sessions, the weight of the own decision is still smaller than that of the given information. However, for the sessions considered separately the picture is less clear. For three sessions (3,5 and 6), the weight of the own decision is slightly larger than that of the given information, for two (2 and 4) it is smaller. Hence we have no signi…cant e¤ect contrary to a false consensus e¤ect in this case, but also no hint in line with the false consensus e¤ect.

On the other hand, it can be argued that the items that are not payo¤ relevant can be included in the analysis since the subjects do not have to estimate what other subjects do but what they state in the experiment. Hence, if subjects do not answer according to their real preferences, they should also take into account the fact that other subjects might do the same. Consequently, the task is essentially the same except that, apart from the possible projection of own preferences, the possible projection of own honesty also matters. Our interest is not primarily in the real preferences of the subjects but in their beliefs about what others state in the same situation. We consider this a valid argument for the inclusion of the not payo¤ relevant items in the analysis.

For all sessions the own decision clearly matters ($\bar{} > 0$) and hence the regression supports the existence of a consensus e¤ect in our experiment. This e¤ect is more clearly established than the underweighting of the own decision. Subjects more consistently weight their own decision positively than that they underweight it.

3.2 Analysis of Items without Information

The average estimates given for the items where we did not provide any information about other subjects' Decision 1 can be found in Table 5. The columns show from left to right a key to the content of the item for Decision 1, the number of A-subjects, their average estimate of A choices, the number of B-subjects, their average estimate of A choices and the di¤erence between the average estimate of A choices given by the A-subjects and by the B-subjects.

3.2.1 The Consensus E¤ect

For all items, the average estimate of A choices is higher for the A-subjects than for the B-subjects. Thus there is a consensus e¤ect at the aggregate level. The same holds for …ve items ("Public Good", "Tax Evasion", "Dona-

item	A #	A ;	B #	B ;	A ; - B ;
Payo¤ System	48	6.8	32	5	1.8
Public Good	58	8.5	22	3.7	4.8
Tax Evasion	42	7.2	38	4.4	2.8
Writing	29	5.7	51	4.2	1.5
Donation ai	41	7.6	39	3.8	3.8
Elevator	57	7.3	23	4.9	2.4
Exam	50	8.1	30	4.4	3.7
Stamps	57	9.1	23	6	3.1

Table 5: Test for consensus e¤ect (own decision not considered as information)

tion ai", "Exam", "Stamps") for each of the …ve sessions. Hence, by using a binomial test as in the preceeding sections, the hypothesis that the estimates are not biased towards the own choice can be rejected at the 5%-level. For the remaining three items, average estimates given by the A-subjects are higher in four out of …ve sessions and hence also clearly in line with a consensus e¤ect, though not signi…cantly. Thus we replicate the results of conventional research on the false consensus e¤ect. To a large extent subjects use the own decision as guideline when estimating other subjects' choices.

3.2.2 The False Consensus E¤ect

Concerning the falsity of the consensus e¤ect when no information is given, our results are di¢cult to interpret. The de…nition we gave above for the (truly) false consensus e¤ect requires that the weight assigned to the own decision can be compared to the weight given to any other subject's decision. Hence, in the absence of information concerning other subjects' decisions, the falsity of a present consensus e¤ect cannot be decided.

An alternative way is to compare the given estimates to those that would be optimal estimates based on Bayesian updating of a reasonable prior belief. Since it is di¢cult to justify a particular prior belief, usually a uniform prior belief is chosen (see Dawes, 1989). Then a false consensus would be considered present if more consensus is estimated than would result from this procedure. Bayesian updating of a uniform prior belief using a sample of size one results in a posterior belief with an expected value of $\frac{2}{3}$ for the own decision (see Dawes 1989). For a group of eleven, this means that the expected value of A choices of the posterior belief is $\frac{2}{3} ¢ 11 = 7:33$ for an own choice of A and $\frac{1}{3} ¢ 11 = 3:67$ for an own choice of B. Since subjects had to give integer estimates, 7 or 4 would have been as close as possible to the expected

values and hence a false consensus could be considered present if A-subjects on average expect more than 7 A choices and B-subjects on average expect less than 4 A choices. According to this criterion, there would have been a false consensus e¤ect for two of the eight items.

However, the assumption of a uniform prior belief (as that of any particular prior belief) is of course completely arbitrary. On the other hand the assumption of a unimodal prior belief with probabilities decreasing in the distance to the mode is justi…ed since anything else would not appear to be a very reasonable form of belief (for example it does not seem plausible to believe that both four and six A choices are more probable than …ve). Such a prior belief would have a smaller variance than a uniform prior belief and therefore the expected value would be more robust to updating. Hence Bayesian updating of such a prior belief would result in a lower di¤erence between expected values of the posterior beliefs of A-subjects and B-subjects. Thus, if we relax the assumption of a uniform prior belief and instead assume a unimodal prior belief, a false consensus e¤ect could be considered present if the observed di¤erence is larger than $\frac{1}{3} \cdot 11 = 3{:}67$: Three items would ful…ll this criterion, two of them just marginally.

Several problems remain with this approach. First of all, it is not necessarily a sign of overweighting of the own decision if too extreme updating is observed. It could simply be that subjects give too much weight to any information about a single choice and this would thus simply indicate an inability to carry out Bayesian updating.[6] Second, there is no reason to give the expected value as estimate, at least with our incentive system. To maximize the expected payo¤, subjects have to maximize the probability of an exact hit and should hence give the mode of the posterior belief as the

[6] In awareness of this possibility Alicke and Largo (1995) designed an experiment to compare the weight of the own decision with information about other subjects. Their design is similar to ours in that they also provide information about four other subjects. Thus we agree with them that a false consensus e¤ect can only be investigated in the presence of information about other subjects.

Alicke and Largo …nd a (truly) false consensus e¤ect, but, in contrast to our study, their items were performance tests. Conceptionally, this is very di¤erent, since the background information concerning performance tests in a familiar area (psychology students had to take a test in personality assessment) is much broader than that of preferences regarding unfamiliar choices. Usually subjects have some reasonable belief about their relative performance and thus information about their success is much more informative than information about the success of a randomly chosen subject. Hence, in this context, it can be perfectly rational to assign greater weight to the own success or failure than to that of another subject.

Alicke and Largo also have to be criticized because their subjects have systematically been deceived and because no incentives for the true revelation of estimates have been provided.

estimate. Under the assumption of a uniform prior belief the mode of the posterior belief is either all A or no A, depending on the own decision. Hence, under these conditions, no overweighting would be possible. And even if an incentive scheme is chosen so that it is optimal to choose the expected value of the posterior belief as an estimate, it will probably not be simple enough to be easily understood by subjects.

If we use the approach to compare estimates with the result of Bayesian updating of a uniform prior belief to individual data, we observe more overweighting (316 choices) than underweighting (219 choices) or perfect updating (105 choices). This could be interpreted as an indication for the presence of a false consensus e¤ect but all the above criticism applies.

The fact that the four items from the category "morally relevant" are among those …ve items with the strongest consensus e¤ect might be considered as support for the hypothesis that a possible false consensus e¤ect has an ego-defensive function. However, this e¤ect is very small for the items from the same category for which we provided information. The consensus e¤ect for the items which are payo¤-relevant is on average stronger than that of those which are not. This contradicts the hypothesis that the consensus e¤ect is, in fact, an adaptation e¤ect which would imply that it would decrease if incentives for the true revelation of preferences are provided. The results for the items with information are along the same line in this respect.

3.3 Additional Results

Two interesting results turned up as side-e¤ects. First, there were surprisingly large contributions to the public good (58 out of 80 subjects contributed), given that this was a one-shot game (although it cannot be excluded that subjects might have expected similar items to follow and thus did not perceive this decision as a one-shot game). Second, the two items "Donation Chapel" and "Donation ai", which were conceptually very similar and thus the respective Decisions 1 were expected to be positively correlated, were indeed slightly negatively correlated (correlation coe¢cient ¡ 0:032).

There is no notable gender di¤erence with respect to either the consensus or false consensus e¤ects. Although the tendency to underweight the own decision is stronger for economics and business students than for the other subjects, this does not allow any conclusion to be made since the data basis for the other subjects was too small. For these subjects, too, there is on average no false consensus, i.e. no overweighting of the own decision, just less underweighting.

3.4 Questionnaires

Although we are careful in interpreting the results of the questionnaires from a general perspective since no incentives are provided for revealing the true beliefs, we still believe that they provide interesting insights into the cognitive processes which caused the results we observed.

The most interesting statements were given by two subjects stating explicitly that they had information not just about four subjects' decisions but about …ve, including themselves, and that they based their estimates on this information. Whereas some subjects decided to discard the information given since they did not consider it as representative, the majority of descriptions of the decision process is well in line with the results we observed. Most subjects stated that, based on the information given, they formed their choice by calculating the corresponding numbers for eleven subjects and then adapting, using either their own choice or general expectations, or both. This process of using their own decision simply as information to adapt the estimates formed on the basis of the information given may have led to an underweighting of the own decision. Subjects obviously tried to be rational but most of them were uncertain about how to weight the own decision, although they were aware that they should consider it somehow. Concerning the rounds without information, the majority of subjects stated that they based their expectations on their own decision, which is clearly in line with the observed consensus e¤ect for this condition.

A surprising result was that subjects on average dramatically overestimated their hit rate. The average estimated hit rate was 28:9%, whereas the average actual hit rate was 11:56%. This is only slightly above the expected hit rate of pure guessing which would have been $\frac{1}{12}$ ¼ 8:33%; a remarkable result on its own (but with the positive consequence for the experimenters of keeping costs much lower than expected). Even more remarkable is the fact that, although the subjects used the provided information in a reasonable way, the actual hit rate for the items without information was slightly better than for the items with information (77 vs. 71 out of 640 possible hits).

4 Concluding Remarks

Our study was designed to provide an easy test for a (truly) false consensus e¤ect and to overcome the shortcomings of previous studies by providing clear monetary incentives for the revelation of true beliefs.

It indicates that, given both these incentives and representative information, although subjects show a consensus e¤ect, they show no false consensus

e¤ect. They signi…cantly underweight, not overweight, their own decision compared to the representative information we provided. This primary result has qualitatively been replicated by A. Ortmann and R. Hertwig who used an adaptation of our design in an experiment at Bowdoin College in December 1998.

Thus the false consensus e¤ect might not be very relevant for economic applications. Plainly, it could even be a result of bad de…nition. However, this conclusion might be premature. It might hold for the type of information we provide, namely explicit information without uncertainty. We do not know what the results would be like if the information were implicit, for example if subjects observed the behavior of other subjects in repeated games, but did not receive a list of others' decisions. In this case, it might happen that subjects overweight their own decisions because of problems in keeping track of others' behavior. This situation may be more relevant to economics since, in reality, people often have the opportunity to observe other people's behavior but do not receive a list of what they did. Certainly this is an area for future research.

O¤erman et al. (1996) provided such implicit information and did not …nd a false consensus e¤ect. They conducted a public good experiment and did not …nd any systematic di¤erence between players and spectators for whom the own decision of the player becomes just another piece of information. Hence, the own decision is not weighted more than the information about other subjects. O¤erman et al. did not even …nd a consensus e¤ect in the sense that cooperators and individualists did not di¤er in their estimates of cooperation rates. However, this result has to be treated with care, since they did not control for the information the subjects received. Given their …nite subject set in each session, cooperators on average met fewer cooperators than individualists did. Hence the equality of the estimates of cooperation rates means that, if subjects used the information at all, they also gave some weight to their own decisions. So there is a consensus e¤ect. But again, this is not a false consensus e¤ect since, including themselves, cooperators on average met more cooperators than individualists did (since they met themselves for sure but only a subgroup of the other subjects in their session). So even if they gave the same weight to their own decision as to that of other subjects, they should have estimated a higher cooperation rate. The results of O¤erman et al. are in line with ours in the sense that there is a consensus e¤ect, but subjects underweight rather than overweight their own choice. What is somewhat surprising in this context is that in our experiment a public good game had the strongest consensus e¤ect among the items for which we did not provide any information.

Concerning the rounds where we did not provide any information, there

are several possible criteria to decide whether the consensus e¤ect we …nd is a false consensus e¤ect. However, we believe that the best conclusion we can draw is that the criticism that can be applied to any one of these criteria is much more convincing then the possible conclusions which it allows. Unless someone comes up with an ingenious idea about how to measure prior beliefs, examining the false consensus e¤ect in the absence of other information does not seem to be a very promising idea. As noted in the introduction, we believe that it does not even make much sense to de…ne a false consensus e¤ect without reference to other information.

References

[1] Abbink, K. and Sadrieh, A., 1995. "RatImage - Research Assistance Toolbox for Computer-Aided Human Behavior Experiments"; University of Bonn, Discussion Paper No. B-325.

[2] Alicke, M. D. and Largo, E., 1995. "The Role of the Self in the False Consensus E¤ect"; Journal of Experimental Social Psychology 31, 28-47.

[3] Camerer, C. 1996. "Rules for Experimenting in Psychology and Economics, and Why They Di¤er"; in: W. Albers et al. (Eds.): Understanding Strategic Interaction – Essays in Honor of Reinhard Selten, Heidelberg: Springer.

[4] Dawes, R. M., 1989. "Statistical Criteria for Establishing a Truly False Consensus E¤ect"; Journal of Experimental Social Psychology 25, 1-17.

[5] Dawes, R. M., 1990. "The Potential Nonfalsity of the False Consensus E¤ect"; in: R. M. Hogarth (Ed.), Insights in Decision Making: A Tribute to Hillel J. Einhorn, Chicago: University of Chicago Press.

[6] Engelmann, Dirk, 1998. "Survival of Trustworthiness as a Consequence of a False Consensus E¤ect"; mimeo, Humboldt-Universität zu Berlin.

[7] Jacobsen, E. and Sadrieh, A., 1996. "Experimental Proof for the Motivational Importance of Reciprocity"; University of Bonn, Discussion Paper No. B-386.

[8] Krueger, J. and Clement, R. W., 1994. "The Truly False Consensus Effect: An Ineradicable and Egocentric Bias in Social Perception"; Journal of Personality and Social Psychology 67, 596-610.

[9] Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D. Story, J. E., and Venderklok, M., 1985. "The False Consensus E¤ect: A Meta–Analysis of 115 Hypothesis Tests"; Journal of Experimental Social Psychology, 21, 263-83.

[10] O¤erman, T., Sonnemans, J., and Schram, A., 1996. "Value Orientations, Expectations and Voluntary Contributions in Public Goods"; The Economic Journal 106, 817-45.

[11] Ross, L., Greene, D., and House, P., 1977. "The 'False Consensus E¤ect': An Egocentric Bias in Social Perception and Attribution Processes"; Journal of Experimental Social Psychology 13, 279-301.

[12] Selten, R. and Ockenfels, A., 1998. "An Experimental Solidarity Game"; Journal of Economic Behavior and Organization 34, 517-39.

[13] Sherman, S. J., Presson, C. C., and Chassin, L., 1984. "Mechanisms Underlying the False Consensus E¤ect: The Special Role of Threats to the Self"; Personality and Social Psychology Bulletin 10, 127-38.

A Experimental Protocol

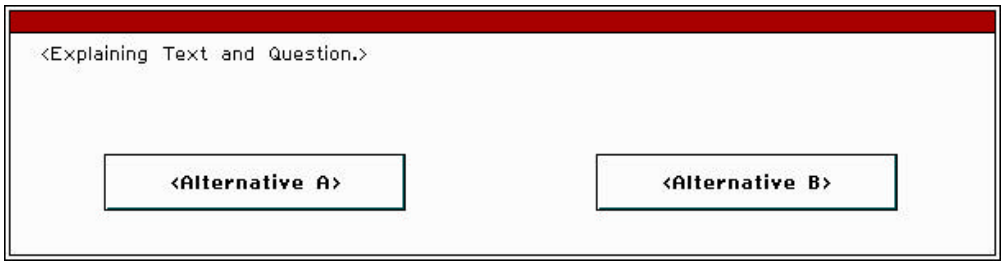
The experiment was conducted at Humboldt-Universität zu Berlin. Subjects where invited to register for an experiment through posters and announcements in classes. When subjects came to the computer pool arranged they were randomly placed at separated, isolated computer terminals. They received written instructions (see Appendix B) for which they could ask for clari…cation. They then played the sequence of Decide, Receive information, and Estimate for 16 di¤erent items (see Appendix C) without any feedback about their hit rate. After this they had to …ll in a questionnaire before getting feedback about their hit rate and payo¤. Finally, they were paid according to their number of hits and their decisions about donations, form and time of payo¤.

B Translated Instructions

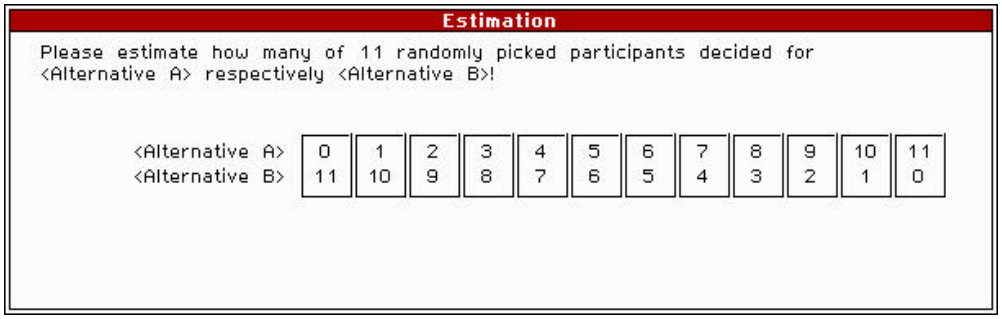
Welcome to our experiment. Please read the following instructions carefully. If you have any di¢culty in understanding, do not ask for help out loud, just raise your hand. We will then answer your questions privately. Please stop communicating with other participants from now on.

You will get a showup fee of DM 10.- for your participation. You can gain or lose additional money during the experiment. You will end up with a positive amount in any case.

The experiment consists of 16 rounds. In each of these rounds you have to make two decisions. First you have to choose between two alternatives (Decision 1). Simply click the relevant button to make your choice.



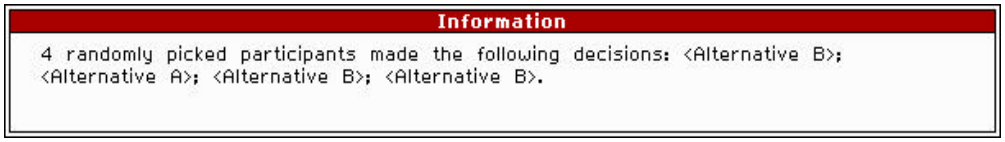
Next your task is to estimate how eleven other subjects decided in Decision 1 (Decision 2). Those eleven subjects are randomly picked from the participants, independently for every participant and round. You will not know who these people are. Please choose a distribution by clicking the relevant …eld. The upper number in each …eld means the number of subjects who have chosen alternative A. The lower number means the number of subjects who have chosen Alternative B.



For example, if you expect that two out of eleven subjects have chosen hAlternative Ai and nine have chosen hAlternative Bi then click the …eld:

2
9

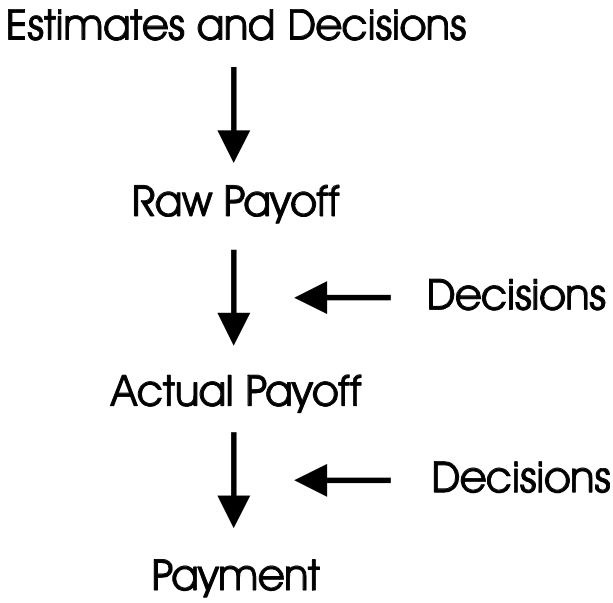
In some rounds you will receive some information before making Decision 2. You will be informed about how four other participants decided in Decision 1. (You do not have to estimate Decision 1 for any of them.)



These four subjects will also be randomly picked independently for every participant and round. Your own decision will not be included. You will not know about whom you received information. Other participants might receive information about your decision but in this case they will not know you were the decision maker.

Your actual payo¤ will be determined as follows: The number of correct estimates in Decision 2 determines your raw payo¤. It consists of DM 10.- showup fee and DM 4.- for every correct estimate (hit). You will not receive any information about your hit rate or your current raw payo¤ during the experiment.

Some of the Decisions 1 have an e¤ect on the size and form of your raw payo¤. There are some transfer decisions which govern how your raw payo¤ is converted into your actual payo¤. As well, there are some decisions which govern the form of your payment. Hence, your payo¤ results according to the following scheme. This may look complicated at …rst, but it will become clear during the experiment.



Please keep the sheet with your subject-id. Without it you won't receive any payment.

C Items

Fruit: What do you eat more frequently, Oranges or Bananas?

Oranges Bananas

Free-riding: Do you occasionally free-ride on public transport?

No Yes

Payo¤ System: You can change the payo¤ system (up to now DM 10.- showup fee and DM 4.- per hit) and receive DM 5.- showup fee and DM 6.- per hit. Do you want to change?

Keep current system Change

Public Good: You can invest DM 1.- of your payo¤ in a shared fund. For each DM invested every participant will receive DM 0.25.

I invest DM 1.- I do not invest

Tax Evasion: Suppose you have a new washing machine delivered. The delivery man o¤ers not to bill you for the sales tax if you go without a receipt. Do you accept this o¤er of tax evasion?

I accept I do not accept

Rich-Poor: If you are the one with the highest raw payo¤, will you give away DM 2.- to the one with the lowest raw payo¤?
I give away DM 2.- I do not give away anything

Lottery: You can take part in a lottery. Your payo¤ is then no longer determined by the hits in all rounds but we will randomly draw one round which is paid out 16 times. The showup fee remains the same.
I take part in the lottery I do not take part

Writing: What do you use more frequently, pencil or pen?
Pencil Pen

Donation ai: You can donate DM 1.- of your raw payo¤ to amnesty international. For every DM donated the experimenters will donate an additional DM 1.-. This donation will be deducted from your raw payo¤.
I do not donate I donate DM 1.-

Time Preference: You can receive your payo¤ immediately after the experiment or you can get your payo¤ plus 10% interest after one month. Do you want your payo¤ now or later?
Now Later

Elevator: If there is an elevator, do you use it for going to the third ‡oor or do you take the stairs?
Stairs Elevator

Exam: Do you allow neighbors who are unknown to you to copy from you in exams?
No Yes

Months: Which month do you prefer for holidays, March or October?
October March

Stamps: You can receive your payo¤ either in cash or in stamps. If you choose stamps you receive one stamp worth DM 1.10 for every DM of your payo¤. We round up to entire stamps. What do you prefer?
Cash Stamps

Donation Chapel: You can donate DM 1.- of your raw payo¤ to the reconstruction of the Heiliggeist Chapel. For every DM donated the experimenters will donate an additional DM 1.-. This donation will be

deducted from your raw payo¤.
I donate DM 1.- I do not donate

Waiter: Do you usually correct the waiter in a restaurant who has miscalculated a bill in your favor?
Yes No