UC Santa Barbara
Departmental Working Papers

Title
Delegating to a Powerless Intermediary: Does It Reduce Punishment?

Permalink
https://escholarship.org/uc/item/0119d201

Authors
Grossman, Zachary
Oexl, Regine

Publication Date
2011-08-30

eScholarship.org Powered by the California Digital Library
University of California

Delegating to a Powerless Intermediary: Does it Reduce Punishment?*

Regine Oexl†, Zachary Grossman‡

August 30, 2011

Abstract

Beyond the classical reasons of efficiency, commitment, the distribution of information, or incentive provision, a person may also delegate decision rights so as to avoid blame for an unpopular or immoral decision. We show that by delegating to an intermediary, a dictator facing an allocation decision can effectively shift moral responsibility onto the delegee *even when doing so necessarily eliminates the possibility of a fair outcome*. Dictators who choose selfishly via an intermediary are punished less and earn greater profits than those who directly choose a selfish outcome, while the intermediary is punished more.

JEL: C91, D23, D63, D86

Keywords: intermediation, delegation, punishment, responsibility, attribution, blame shifting, experimental economics, behavioral economics

1 Introduction

Why do managers hire consultants to announce and implement layoffs, as well as outsource final production decisions?[1] Why might a company like Merck sell the patent to a cancer drug to another company (Ovation), likely anticipating that this will significantly inflate the price, instead of doing so directly?[2] While firms hire agents for reasons of efficiency, commitment or incentive

*We wish to thank to Antonio Niocolò for helpful discussions. Furthermore we are grateful for valuable comments to Marco Piovesan, Gary Charness, and the seminar participants at the International Meeting on Experimental and Behavioral Economics (IMEBE) 2011, Barcelona, 4th Maastricht Behavioral and Experimental Economics Symposium (M-BEES 2011), Maastricht and at the Work in Progress Seminar, Dipartimento di Economia Politica, Università degli Studi di Milano-Bicocca, May 2011. This work is part of the PhD thesis of Regine Oexl. She thanks to the "Cassa di Risparmio di Padova e Rovigo" for financial assistance during her PhD. The usual disclaimer applies.

†Dipartimento di Scienze Economiche "M. Fanno", Università di Padova, regine.oexl@unipd.it

‡Department of Economics, University of California, Santa Barbara, grossman@econ.ucsb.edu

[1] In the 2009 movie "Up In The Air", George Clooney portrays such a professional corporate downsizer. Hamman et al. (2010) list other sources of illustrative examples, including "Letting a Stranger Do the Firing", Paul Brown 2007, The New York Times Nov. 10 http://www.nytimes.com/2007/11/10/business/10offline.html and O'Rourke (1997), who considers a case study of Nike's working standards in outsourced plants as well as the accounting firm's labor and environmental audit.

[2] Coffman (2011) and Paharia et al. (2009) provide the background on Merck. Merck sold the patent of a cancer drug to a small company, Ovation, that subsequently raised the price of this drug by more than a factor of ten. Since Merck had sold the rights for the drug, the public generally did not perceive it to be responsible for this increase in price. However, after paying a high enough purchase price, Ovation might have been compelled to increase the drug price to avoid a loss. Thus, by demanding a high enough price for the rights to the cancer drug, Merck might effectively limit Ovation's strategy space to 'unfair' price increases. Of course, this is conditional on Ovation choosing to buy the patent. An area ripe for future research is an agent's decision of whether or not to accept the intermediary role.

provision (Aghion and Tirole, 1997, Bolton and Dewatripont, 2005, Schelling, 1960), responsibility-shirking and blame-shifting provide an additional rationale for delegating decision rights. Previous studies (Bartling and Fischbacher, 2011, Coffman, 2011, Fershtman and Gneezy, 2001, Hamman, Loewenstein, and Weber, 2010) have found that players in experimental dictator games may avoid censure and costly punishment by delegating the allocation decision to an intermediary. However, because those studies provide the intermediary either with options with differing degrees of fairness, or with no choice whatsoever, it is not clear whether the willful choice of the intermediary of a selfish outcome over a less-selfish outcome is necessary for blame shifting, or whether the mere presence of a nominal intermediary is sufficient, regardless of her power to influence the fairness of the outcome.

We conducted an experiment, detailed in Section 2, in which a dictator may choose from one of three allocations of \$20, including an equal split, among her four-person group. She may also delegate her decision to an intermediary, but only in such a way that limits the intermediary's choice set to the two unfair allocations, that differ merely in which of the two passive group members (recipients) are hurt most.[3] For a cost, the two receivers can reduce the earnings of some or all of the other players in the group, conditioning their deductions on the choices of the dictator and intermediary. Though the experiment instructions avoid all language pertaining to punishment, we follow Bartling and Fischbacher (2011) (hereinafter, BF) by interpreting the deductions chosen by a recipient as a form of punishment, reflecting how he attributes responsibility to the other players for the unfair outcome.

Section 3 presents the results. We find that delegating is profit maximizing. Thus, even though the intermediary, if called upon, is powerless to implement the fair outcome, the dictator can effectively shift responsibility when delegating. Given an unfair outcome, dictators experience smaller deductions when they use the intermediary instead of choosing directly, and the intermediaries in turn experience larger deductions.

These findings complement and extend previous results on intermediation and blame-shifting. Paharia et al. (2009) find that when asked to rate unethical behavior in hypothetical scenarios, subjects judge acts carried out through an intermediary more leniently, though only when shown them one-by-one. Our design most closely resembles the delegation and punishment (D & P) condition of BF, who find find that the dictator's punishment given an unfair allocation is greatly reduced when the dictator passes the choice to the intermediary. While in the BF design, the act of delegating forwards the whole action set to intermediary, in our design the intermediary can only choose between the unfair options. Thus, it is transparent that a dictator who delegates intends an unfair allocation to be obtained and that the intermediary has no say in whether or not the fair outcome was chosen. As in the *asymmetric* condition of BF, in which the dictator cannot choose the unfair outcome directly, delegation can be considered the least kind action available to the dictator (along with choosing an unfair allocation directly). This makes it all the more striking that the dictator can effectively shift blame to the intermediary by delegating and suggests that intentions alone cannot explain the punishment behavior.

Instead, our findings, like those of Coffman (2011), suggest that the mere fact that the dictator does not *directly* determine the final allocation is sufficient for blame-shifting. In Coffman's design, a dictator can involve the intermediary in a way that leaves her with room to choose between allocations with varying degrees of fairness, or in the extreme, she may leave the intermediary

[3] In the following, we will refer to the dictator as "she", while the intermediators and the recipients will be referred to by "he".

with no choice but the least fair allocation. In our design, any involvement of the intermediary necessarily means that an unfair outcome will result, but delegating always leaves the intermediary with a non-degenerate choice. The fact that dictators in our experiment are punished less for an unfair outcome that they obtained indirectly echoes Coffman's findings, but, unlike Coffman, we observe that delegating shifts blame *onto* the intermediary, rather than merely reducing the blame placed on the dictator. Thus, even when the intermediary can not in any way be viewed as complicit in the the unfair behavior, she may be punished for it.

Not only did the participants in the intermediary role lack the ability to enforce a fair outcome, instead of choosing or contracting on their role as an intermediary, it was assigned to them. The fact that intermediaries are still punished for the unfair outcome raises the question of why an agent would agree to perform a blame-worthy task for a principal looking to avoid punishment. The stated beliefs of our participants suggest an explanation. While participants' beliefs are qualitatively in tune with the observed punishment patterns, the dictators and particularly the intermediaries underestimate the extent to which blame is shifted onto the intermediary.

Given the reduction in punishment, a money-maximizing dictator with rational expectations has the incentive to delegate. Furthermore, by delegating a dictator might also avoid the uncomfortable decision (see Dana et al., 2006, 2007, Grossman, 2010, Lazear et al., 2009) of having to decide whom to harm most, while avoiding the cost of choosing the equal split. However, we find that roughly as many dictators directly choose an unfair allocation as use the intermediary. Surprisingly, the degree to which the dictator perceives intermediation as a way of avoiding punishment does not significantly increase the likelihood that the dictator will delegate. We conclude that delegating one's harmful action to someone else might "feel wrong" or violate a personal rule, imposing a psychological cost.

We conclude in Section 4 with a discussion of how outcome- and intentions-based theories of social preferences have difficulty explaining our results. Our findings challenge the responsibility measure proposed by Bartling and Fischbacher, because we find the intermediary to be punished even when her action cannot possibly increase the likelihood of an unfair allocation, which is already certain. Instead, our experiment reinforces Coffman's conclusion that recipients are less prone to sanction harmful behavior when the responsible party does not directly interact with the victim. The fact that the dictator can effectively shift blame to the intermediary when her choice so transparently binds that of the intermediary reinforces the importance of delegation as a tool to evade the perception of culpability and raises questions about an agent's decision to take on the intermediary's role.

2 Experimental Design

The experiment was computerized with the software Z-tree (Fischbacher, 2007). We conducted 14 sessions, lasting 30 - 45 minutes each and featuring 41 groups and 164 subjects, in May - June 2010 and January 2011.[4] We used the the online system ORSEE (Greiner, 2003) to randomly recruit participants from the University of Santa Barbara (UCSB) Experimental and Behavioral Economics Laboratory (EBEL) subject pool, largely comprised of UCSB students and staff. Upon arriving at the experiment, participants sat at computer terminals, were given a paper copy of the instructions,

[4]We dropped observations from two subjects who participated in more than one session, in violation of the experiment protocols.

and followed along as the experimenter read them aloud. We randomly assigned participants into four person groups, in which the roles of the dictator (D), intermediator (N), and two recipients (R_1 and R_2) were randomly assigned.

We then gave participants a second packet of written instructions explaining their specific role more in detail. These instructions included exercises designed to verify participants' understanding of the instructions, which the experimenter verified before the the decision-making began. Average payment was \$10.51, including the show-up fee. Full instructions as well as screenshots are presented in the appendix.[5]

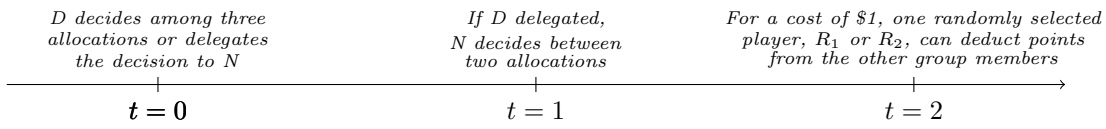


Figure 1: Timeline

Each group played a one-shot dictator game once, with timing as follows (see also Figure 1). First, the dictator chooses between one of three allocations of \$20 between the four players, or she can delegate the decision to the intermediator. The dictator's four choices are presented in Table 1). Allocation c is costly but fair, assigning \$5 to each player, while a and b are two very similar unfair allocations, differing only in which receiver receives less. They both benefit the dictator and the intermediary, allocating each \$9, leaving the remaining \$2 to one of the two recipients. Then, if D chose to delegate the decision, N was required to choose *only* between the two unfair allocations, a and b, with no possibility of choosing the equal allocation.[6] We shall refer to the realized split of the \$20 as an *allocation*, whereas a terminal history of the decision process, including both the allocation and the player choosing it, will be referred to as an *outcome*.

Table 1: The dictator's four choices. The intermediary must choose either a or b if D passes.

Choice	Dollars allocated to			
	D	N	R_1	R_2
a	9	9	0	2
b	9	9	2	0
c	5	5	5	5
d	(Pass to N)			

Finally, the two recipients had the opportunity to punish the other players by reducing their payoff. The recipients could deduct up to a total of \$7 from the payoff of any combination of the other participants, with the restriction that the resulting payoffs must be non-negative. If the recipient deducted a non-zero amount from any of the other players, he paid a flat fee of \$1. We used the strategy method to elicit the punishment choices for both R_1 and R_2; each had to decide how many punishment points to allocate to players D, N, and the other receiver in each of the

[5]The software is available from the authors upon request.

[6]Instead of stating a contingent strategy, the intermediary only made a decision when called upon after the dictator delegated.

five possible scenarios (in a randomized order). At the end of the experiment, either R_1 or R_2 was randomly selected, and the punishment specified by that player's strategy for the realized outcome was implemented.

In the last four sessions, after having completed their task, we elicited additional beliefs of 12 dictators and 12 intermediaries. Participants stated the amount they expected R_1 to deduct from each of the subjects in their group, receiving an extra payment of \$5 if the stated amount was in a range of +/- \$1 of the true average elicited in previous sessions. Having two recipients allows us, like BF, to double the number of punishment observations per group. Importantly, it also allows us to provide the intermediary with a choice between *two* unfair outcomes, instead of only one.

3 Results

3.1 Punishment

First we examine punishment behavior of the recipients. Figure 2 shows the pattern of punishment across outcomes for the dictator and intermediary and illustrates our three main results. Recipients appear to punish unfair behavior, but shift the blame for a delegated unfair outcome from the dictator to the intermediary, making delegating the profit maximizing choice for the dictator.

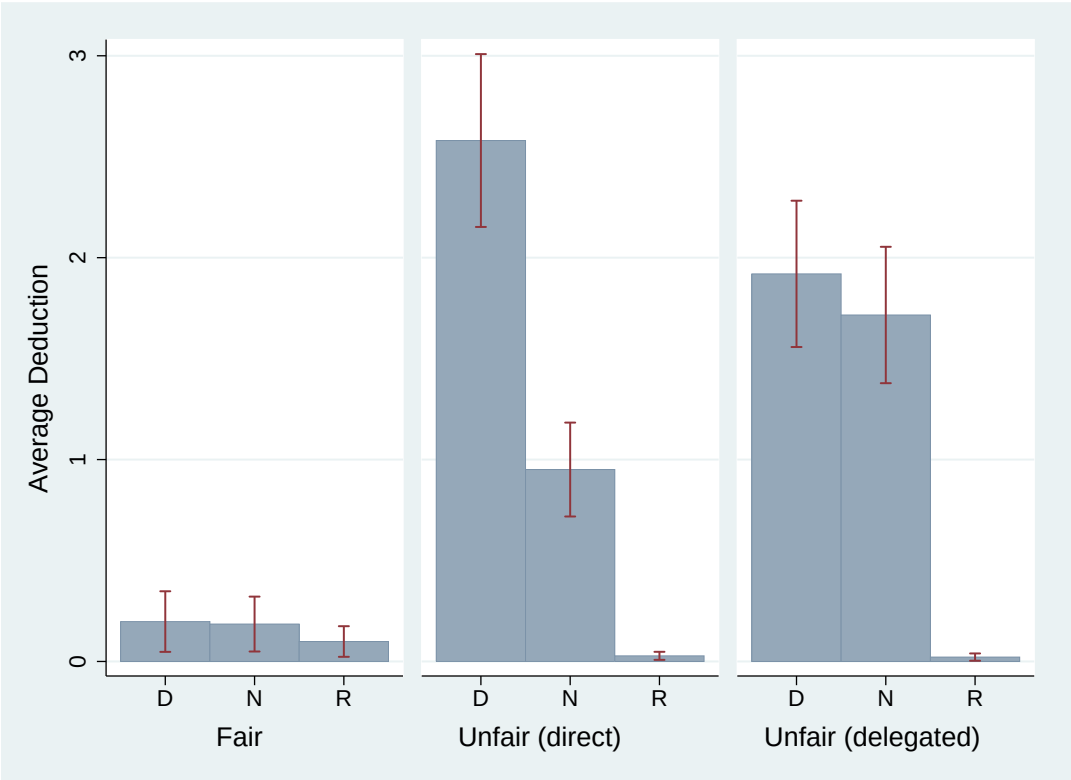


Figure 2: Average deduction, including all recipients, by outcome

In this section we provide support for these observations. The data is summarized in Table 2, which shows the frequency of punishment and the average deduction amounts among recipients who chose non-zero deductions for a given outcome. More than 70% of the 81 recipients chose to

incur the \$1 cost to deduct money from at least one other player for at least one outcome.

Table 2: Rate of non-zero deductions for each outcome by recipient and the average amount deducted from the dictator and recipient.

Outcome	Deduction Rate[a]		Mean Deduction[b] from D by		Mean Deduction[b] from N by	
	R_1	R_2	R_1	R_2	R_1	R_2
a chosen by D	19	21	5.05	4.81	1.63	1.95
b chosen by D	22	24	5.09	4.54	1.59	1.96
c chosen by D	4	3	1.75	3.00	2.00	2.33
a chosen by N	22	20	3.23	3.45	3.40	3.25
b chosen by N	22	25	3.91	3.40	2.64	3.20

[a] The total number of respective R_1 and R_2 recipients was 41 and 40.
[b] Mean is conditional on non-zero deductions for the outcome.

Recall that the two unfair allocations differ in which recipient is harmed the most: allocation a gave R_1 a payoff of \$0, as opposed to the \$2 he obtained from allocation b, with the reverse payoffs for R_2. To begin our analysis, we note that, holding constant whether it was chosen by the intermediary or directly by the dictator, recipients do not appear to distinguish one of these allocations from the other in their punishment decision.

Result 1. *Punishment is not sensitive to which unfair outcome was chosen.*

When the dictator chose an unfair allocation directly, 43 (53%) out of the 81 recipients who were *most harmed* (R_1 for a and R_2 for b) chose to incur the \$1 punishment cost and exactly the same number of recipients who were *least harmed* (R_2 for a and R_1 for b) did. The zero difference in punishment frequency between recipient types is not significant at any level by any hypothesis test. When the dictator delegated, 47 (58%) of the recipients who were most harmed chose to incur the punishment cost, while 42 (52%) of those who were harmed least chose to punish. A two sample t-test cannot reject the hypothesis that these punishment rates are the same ($t = 0.79$, $p = 0.43$).

Examining the mean and distribution of the deductions chosen by those who did punish leads to the same conclusion. For the dictator's punishment, a two-sample t-test with unequal variance cannot reject the hypothesis ($t = 0.48$, $p = 0.63$) that the same amount was deducted on average by the least-harmed (4.95) and most-harmed recipient (4.77) when the dictator chose a or b directly. A two-sample Kolmogorov-Smirnov test cannot reject the hypothesis that the distribution of deductions are equal ($p = 0.90$). When D delegated and N chose, 3.40 was deducted on average by the most-harmed recipient and 3.69 was deducted by the least-harmed. We cannot reject the hypothesis that these deductions are equal in means ($t = 0.53$, $p = 0.60$) or distribution (KS $p = 1.00$).

The same is true for the intermediary's punishment. When D chose directly, the mean deduction from N by the most-harmed recipient, 1.81, is not significantly different ($t = 0.13$, $p = 0.90$) from that made by the least-harmed recipient, 1.77, nor are the distributions significantly different (KS $p = 1.00$). When D delegated, the mean deduction from N by the most-harmed recipient, 3.30, is not significantly different ($t = 0.85$, $p = 0.40$) from that made by the least-harmed recipient, 2.93, while the distributions again are not significantly different (KS $p = 0.98$).

Given that the recipients treat the two unfair allocations identically, for the remaining analysis

we pool the punishment data for two unfair allocations within the two categories: unfair allocations chosen directly and unfair allocations that were the result of a delegated choice. We continue by observing that the recipients' deductions are consistent with punishing unfair or harmful behavior.

Result 2. *The recipients punish the unfair outcomes more than the fair outcome.*

Whereas only 7 recipients (9%) chose to punish the dictator for choosing the equal split, c, no fewer than 40 (49%) chose to punish when one of the non-equal allocations was chosen, whether directly or indirectly, with an average of 54% of the recipients punishing across all unfair outcomes. A two-sample t-test rejects the hypothesis that the punishment rates are equal ($t = 16.28$, $p < 0.01$).

Furthermore, among recipients who did deduct, the mean punishment is higher when either unfair option was chosen compared to when c was chosen.[7] The dictator receives an average punishment of \$2.28 when choosing the fair allocation, while getting punished \$3.49 when delegating and getting deducted \$4.86 when choosing a or b directly. A two-sample t-test with unequal variance rejects the hypothesis that the mean deduction is the same whether the dictator chose c or she chose a or b ($t = 10.83$, $p < 0.01$) as well as the hypothesis that the mean deduction is the same whether she chose fairly or delegated ($t = 3.34$, $p < 0.01$).

The intermediary received an average punishment of \$2.14 when D chose c, an average punishment of \$3.12 when D delegated, and an average punishment of \$1.79 when D chose either a or b directly. While N's punishment for D's direct unfair choice is smaller than for D fair choice, a two-sample t-test with unequal variance can not reject the hypothesis that the mean punishment is equal. Furthermore, as Figure 2 illustrates, the unconditional mean deduction following c of 0.19 is significantly smaller than the 0.95 that followed a direct a or b choice ($t = 7.87$, $p < 0.01$).

Figure 2 also shows that although the dictator's allocation is the same for a or b regardless of how it was chosen, her final payoff is not. Because, given an unfair outcome, delegation reduces the dictator's punishment, she earned the most when she delegated.

Result 3. *Delegation is profit maximizing.*

Whereas D received an average payoff of \$6.00 for choosing either a or b and \$5.67 from choosing c, she averaged \$8.13 when delegating. An unpaired t-test with unequal variance rejects at the 7% level ($t = 1.94$) the hypothesis that D earns the same amount whether choosing an unfair allocation directly or indirectly and a two-sample Kolmogorov-Smirnov test rejects the hypothesis that the distribution of earnings for the two outcomes is the same ($p < 0.01$). Similarly, a two-sample t-test with unequal variance rejects the hypothesis ($t = 2.61$, $p = 0.02$) that D earns the same payoff when delegating as when choosing fairly. Thus, we conclude that delegating is profit maximizing.[8]

Result 4. *While delegation reduces the blame placed on the dictator, overall blame does not diminish: it just shifts toward the intermediary.*

Our final punishment result is the observation that, given the unfair outcome, delegating increases the intermediary's punishment. We reject the hypothesis ($t = 4.74$, $p < 0.01$) that the mean deduction from N is the same whether an unfair allocation is chosen directly (\$1.79) compared to when the choice is delegated (\$3.12). Similarly, we cannot reject ($t = 0.24$, $p = 0.81$) the hypothesis

[7] All the following results are reported conditional on punishing; however, *all* the qualitative results below also hold unconditionally.

[8] Interestingly, we cannot reject the hypothesis that D's punishment for choosing a or b directly renders her net payoff equal to what she would have earned from choosing fairly ($p = 0.705$).

that, even though delegation reduces D's punishment for an unfair outcome, the average *combined* punishment meted out to the dictator and intermediary is the same whether the dictator chooses directly (3.53) or she delegates (3.64).

3.2 Dictator Behavior and Beliefs

Next we look at the dictator's behavior, which is summarized in Figure 3. First, we observe that not all dictators choose the profit-maximizing way to implement a selfish outcome.

Result 5. *A significant number of dictators forgo the intermediary and directly choose an unfair outcome.*

Fifteen (37.5%) out of 40 dictators delegated the decision to the intermediary, while 12 (30%) chose the fair allocation and 13 (32.5%) chose a or b directly. A two-sample test of proportions does not reject the hypothesis ($Z = 0.47$, $p = 0.64$) that the percentage choosing a or b directly and the percentage choosing to delegate are equal. Nor can we reject the hypothesis ($Z = 0.24$, $p = 0.81$) that the percentage of dictators choosing a or b directly and the percentage choosing fairly are equal.

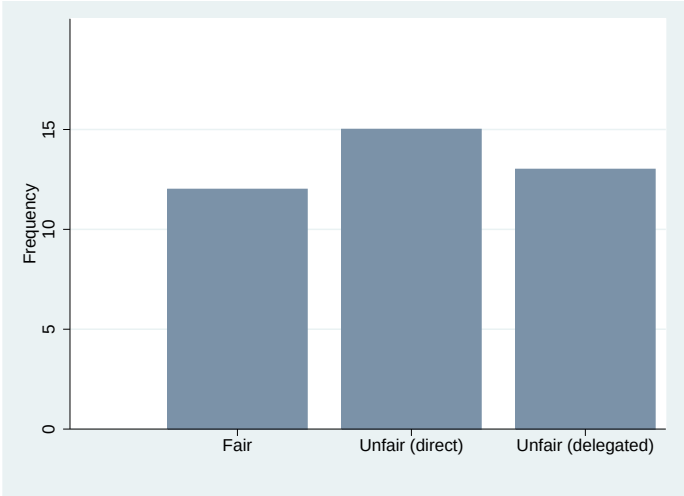


Figure 3: The dictators' choices

While social preferences such as inequity aversion or simple altruism can explain why some dictators might choose the fair allocation, this result raises the question of why any dictator would choose an unfair allocation directly, instead of delegating the choice to the intermediary. To answer this question, we turn to the beliefs expressed by the dictators. First, we note the accuracy of the aggregate beliefs: the dictators correctly predict the qualitative features of the punishment schedule, on average.

Result 6. *The average beliefs of the dictators about the deduction behavior of the receivers are in line with the observed pattern of deductions.*

Table 3 summarizes the mean beliefs of the dictators and intermediaries about the average deductions corresponding to each outcome. On average, dictators expect to be punished the most for choosing a or b directly ($3.04), a low punishment for choosing the equal allocation ($0.75) and an

intermediate amount ($1.42) for delegating. We can reject the hypothesis that the dictators dictators expressed beliefs about the average punishment for choosing unfairly are the same regardless of whether or not she delegated the choice. This is true both of the mean belief ($t = 3.07$, $p < 0.01$), which is $3.04 for the direct choices and $1.42 for delegating, as well as for the distribution of the beliefs (KS $p = 0.04$). The intermediaries expect a similar pattern. We can reject the hypothesis ($t = 2.82$, $p < 0.01$) that the intermediary expects the same average punishment for the dictator regardless of whether she chooses a or b directly ($3.04) or she delegates ($1.33).

Table 3: Mean beliefs of dictators and intermediaries about average deductions

Outcome	Dictator's beliefs about the deduction from:			Intermediary's beliefs about the deduction from:		
	D	N	R^a	D	N	R
a chosen by D	3.33	1.75	0.33	2.83	1.42	0.33
b chosen by D	2.75	1.08	0	3.25	1.33	0
c chosen by D	0.75	0.5	0.33	0.42	0.67	0.58
a chosen by N	1.67	2.67	0.25	1.25	1.5	0.25
a chosen by N	1.17	1.83	0	1.42	2	0

[a] R is defined as whichever recipient is not making the punishment decision.

The beliefs of the both the dictator and the intermediary about the intermediary's punishment conform to the observed pattern to a lesser degree. Although the difference is not statistically significant ($t = 1.58$, $p = 0.12$), the mean belief expressed by dictators about the average amount ($2.25) deducted from N when she delegates exceeds that following her direct choice of an unfair allocation ($1.42). Similarly, the average N believes that the average punishment drops from $1.75 to $1.38 when the dictator chooses an unfair allocation directly instead of delegating, but this difference is not significant ($t = 0.75$, $p = 0.46$).

To what extent, though, do dictators' beliefs explain behavior at the individual level? Though the relatively small number of dictators (12) from whom we elicited beliefs does not permit much rigorous statistical analysis, we do not see any evidence that those who choose a or b directly perceive a smaller punishment reduction for delegating.

Result 7. *The difference in expected deduction between choosing unfairly directly and via the intermediary is not a good predictor of delegating.*

Eight out of twelve dictators for whom we have beliefs data chose an unfair allocation, with two choosing directly and six delegating. The two who choose directly on average perceive that their punishment would be $4.50 *lower* had they delegated, while those who delegated perceived the savings from doing so to be only $2.50. Thus, it is hard to argue from this data that those who failed to delegate did not perceive as much of a monetary advantage from doing so as those who did.

4 Conclusion

We conducted an experiment in which a dictator could choose either an equal split of a $20 endowment among members of her four-person group, or one of two unfair allocations that increased the payoff for her and a second group-member, at the expense of the two remaining group members, or she could allow the second player to choose between the two unfair allocations. One of the passive recipients was allowed to pay a small cost to deduct up to $7 from any of the other participants, contingent on the outcome. We find that the recipients use the deductions to punish unfair behavior, but that dictators can effectively shift the blame for unfair behavior onto the intermediary, even though delegating necessarily leaves the intermediary with no choice but to select an unfair allocation.

Why can the dictator avoid punishment by delegating? Outcome-based theories of social-preferences (Bolton and Ockenfels, 2000, Charness and Rabin, 2002, Fehr and Schmidt, 1999) cannot explain the shift in punishment from the dictator to the intermediary when the intermediary, as opposed to the dictator, chooses an unfair allocation. Nor can intentions-based models of sequential reciprocity, such as Dufwenberg and Kirchsteiger (2004), which regard the dictator's decision to delegate as no less fair than choosing an unfair allocation directly and which do not regard any of intermediary's choices as unkind, given that she is limited only to unfair options.

Though, unlike Coffman (2011), our experiment was not designed with the explicit purpose of eliminating alternative explanations, such as limited reasoning or the mere presence of a third party, our result is entirely consistent with Coffman's finding that the punishment decreases if the dictator avoids interacting directly with the harmed recipient. However, much remains to be understood about the mechanism behind this effect. Our results show that it persists even when the intentions of the dictator and the absence of responsibility of the intermediary are very transparent, which suggests that delegating 'dirty work' can effectively reduce blame, creating a powerful motive for delegation in a wide variety of settings.

On the other hand, like BF, we observed blame actually being shifted *onto* the intermediary, as opposed to merely off of the dictator. This highlights an important limitation on the effectiveness of delegation as a way avoiding the costs of improper behavior. In some cases, in order to make a contract individually rational for an agent with rational expectations, a principal would have to compensate the agent for much of the punishment she herself would have incurred anyway. Further research with voluntary contracting between the principal and agent, as in Hamman et al. (2010), is needed to understand the conditions under which principals can successfully avoid the costs of misbehavior in the marketplace.

Bartling and Fischbacher propose a formal measure of responsibility to explain punishment patterns. According to their measure, a player takes on responsibility for a bad outcome if and only if her action increases the probability that this outcome will result. While their responsibility measure outperforms outcome- or intentions-based social-preference models in predicting punishment behavior in their experiment, it does not pass the stress test to which we subject it in our experiment. In our setting, a delegating dictator increases the probability that an unfair outcome results, while the intermediary does not, for when the choice comes to her an unfair outcome is already guaranteed. Thus, the BF responsibility measure cannot explain why blame is shifted from the dictator to the intermediary when the dictator delegates in our experiment.

Using the intermediary is profit maximizing and may avoid possible psychological costs asso-

ciated with having to choose whom to harm most. So why do roughly a third of the dictators choose an unfair outcome directly? Unlike in the experiment of Bartling and Fischbacher, who find a similar result, there is no risk that the intermediary will choose the fair outcome, leaving the dictator with a low payoff.

While it is possible that these dictators care about the wellbeing of the intermediaries and are willing to reduce their own payoff to avoid lowering that of the intermediary, this seems highly unlikely. These dictators can hardly be described as averse to inequality, having forgone the punishment-minimizing equal split. Furthermore, even if these dictators care about the intermediary to the exclusion of the recipients, perhaps including only the intermediary in her reference group because they both benefit from the unfair allocation, the fact that helping the intermediary increases inequality and yields a material payoff for the dictator that is lower than that of the intermediary renders this explanation even less plausible.[9]

Some dictators may have directly chosen an unfair allocation under the belief that the punishment savings from delegating would be trivial, or that the recipient would punish them *more* for delegating because doing so includes and innocent third party and highlights the dictator's "dubious motive" (Paharia et al., 2009). However, we find no link between propensity to delegate and the perceived difference in punishment between direct versus delegated unfair allocations. A more plausible explanation is that delegating one's harmful action to someone else might "feel wrong" or violate a personal rule, imposing a psychological cost.

Individuals are known to exploit or even create situations in which observers cannot clearly link actions with their consequences, in order to reduce the social consequences of selfish or harmful behavior. We have shown that, through the use of an intermediary, a dictator may escape social sanctions even when her behavior is *transparently* harmful. Furthermore, despite the transparency, she may effectively shift the blame onto the intermediary. It remains to be seen just how transparently limited the role of the intermediary may be yet still permit the dictator to shift the blame.

These findings have important implications, not just for interpersonal or public relations, but also in organizational settings. By delegating an unpopular decision, a manager may limit the impact on employee morale and effort provision, deflecting negative feelings towards a "fall guy" or an outside company. However, we still know little about the functioning of 'directness' and about the mechanics of blame-shifting and the delegator-intermediary relationship. Clearly, the punishment technology and the extent to which the intermediary is even subject to sanctions are important. Our finding of blame-shifting, and that of BF, were obtained when punishment was costly and constrained, while Coffman uses a costless and unconstrained technology, which may account for the fact that he did not observe high levels of intermediary punishment.

To what extent can one continue to shift the blame onto the intermediary over the long run? Surely an agent that is punished regularly in the place of the party for whom she serves as the intermediary would require compensation. A savvy principal might select and agent that cannot be easily or effectively punished, such as when a manager hires an outside consultant to implement an unpopular decision. Those passing judgment might eventually lay blame where it is deserved, which may make it difficult to rely on a blame-shifting strategy too heavily and persistently.

[9] Charness and Rabin (2002) find that people have little willingness to sacrifice their own payoff to help those who have more than them.

References

Philippe Aghion and Jean Tirole. Formal and real authority in organizations. *Journal of Political Economy*, 105(1), 1997.

Björn Bartling and Urs Fischbacher. Shifting the blame: On delegation and responsibility. *Review of Economic Studies*, 2011.

Gary E. Bolton and Axel Ockenfels. Erc: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193, 2000.

Patrick Bolton and Mathias Dewatripont. *Contract Theory*, volume 1 of *MIT Press Books*. The MIT Press, June 2005.

Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117, 2002.

Lucas C. Coffman. Intermediation reduces punishment (and reward). *American Economic Journal: Microeconomics*, November 2011.

Jason Dana, Daylian M. Cain, and Robyn M. Dawes. What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100 (2):193–201, July 2006.

Jason Dana, Roberto Weber, and Jason Kuang. Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1):67–80, October 2007.

Martin Dufwenberg and Georg Kirchsteiger. A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268–298, May 2004.

Ernst Fehr and Klaus M. Schmidt. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868, 1999.

Chaim Fershtman and Uri Gneezy. Strategic delegation: An experiment. *The RAND Journal of Economics*, 32(2):352–368, 2001.

Urs Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 2007.

Ben Greiner. An Online Recruitment System for Economic Experiments. *Forschung und wissenschaftliches Rechnen*, 63:79–93, 2003.

Zachary Grossman. Strategic ignorance and the robustness of social preferences. University of California at Santa Barbara, Economics Working Paper Series 1469100, Department of Economics, UC Santa Barbara, August 2010. URL http://ideas.repec.org/p/cdl/ucsbec/1469100.html.

John R. Hamman, George Loewenstein, and Roberto A. Weber. Self-interest through delegation: An additional rationale for the principal-agent relationship. *American Economic Review*, 100(4): 1826–46, September 2010.

Edward Lazear, Ulrike Malmendier, and Roberto Weber. Sorting and social preferences. NBER Working Papers 12041, National Bureau of Economic Research, Inc, February 2009.

Dara O'Rourke. Smoke from a hired gun: A critique of nike's labor and environmental auditing in vietnam as performed by ernst & young. *report published by the Transnational Resource and Action Center: San Francisco*, November 1997. URL www.corpwatch.org/trac/nike/ernst/.

Neeru Paharia, Karim S. Kassam, Joshua D. Greene, and Max H. Bazerman. Dirty work, clean hands: The moral psychology of indirect agency. *Organizational Behavior and Human Decision Processes*, 109(2):134 – 141, 2009.

Thomas Schelling. *The Strategy of Conflict*. Harvard University Press, Cambridge, MA, 1960.

A Instructions

The following instructions were given both in written and oral form, while in appendix A.1 - A.3 the instructions are presented which were read by each participant. The examples and exercises participants had to solve are presented in appendix A.3.3.

(Oral) Welcome and thank you for participating in this decision-making experiment. You will be paid privately in cash at the end of the experiment, which will last around half an hour. Research foundations have provided the funds. You will make a few decisions that will affect your payoff and possibly the payoffs of other participants. Other participants will simultaneously be making choices that may affect your payoff. Please pay careful attention to the instructions as a considerable amount of money is at stake. You are guaranteed a minimum payment of $5, and may earn as much as $15.

Your participation in this session and any information about your earnings will be kept strictly confidential. Your payment-receipt and consent form are the only places in which your name or perm number are recorded. You will never be asked to reveal your identity to anyone during the course of the experiment. In order to keep decisions private, please do not reveal your choices to any other participant. You will complete one task. Your earnings will be calculated based on your decision and the decision of other subjects, and at the end of the experiment you will be paid that amount plus a $6 show-up fee. If you have any questions during the experiment, please raise your hand and wait for assistance. Please note that for each screen, once you click OK you cannot go back to the previous screen. Please make sure you have read and understand everything completely before you move on.

In this experiment, you will be anonymously grouped together with three other people, so that your decision may affect the payoffs of these three, just as the decisions of the other people in your group may affect your payoffs. You will not know the identity of the other people, and the other people will not know your identity. Each group will consist of four kinds of players, one participant A, one participant B, one participant C and one participant D. Decisions will be made sequentially, in alphabetical order. Participant A starts and can decide how to divide 20 dollars between the four participants. Participant A can choose between four options:

- Distribution 1: Participant A and participant B each receive 9 dollars, participant C receives 2 dollars and participant D receives 0 dollars (this is: 9,9,2,0).
- Distribution 2: Participant A and participant B each receive 9 dollars, participant C receives 0 dollars and participant D receives 2 dollars (this is: 9,9,0,2).
- Distribution 3: Each participant receives 5 dollars (this is: 5,5,5,5).

Participant A can choose between these three distributions, or can pass the decision to participant B. Then, participant B can decide among the first two distributions:

- Distribution 1: Participant A and participant B each receive 9 dollars, participant C receives 2 dollars and participant D receives 0 dollars (this is: 9,9,2,0).
- Distribution 2: Participant A and participant B each receive 9 dollars, participant C receives 0 dollars and participant D receives 2 dollars (this is: 9,9,0,2).

If A delegates the decision to participant B, A cannot take any further decision. Once A, or, in case A delegates, B has made a choice about the allocation of the 20 dollars, participant C and participant D are informed about

- if A has delegated the decision or not
- what distribution was chosen.

Then, either participant C or participant D is randomly chosen. This participant has the possibility to deduct dollars from A, B and the other participant (either D or C), at the cost of one dollar. The randomly chosen player can deduct a maximum of 7 dollars, but can also deduct less. The player deducting points cannot deduct more point from a participant than that participant has earned through the chosen allocation.

Next you will be assigned a role. You will get further instructions on paper, which explain how you will make the decisions for your specific role. There are some examples, and seven short exercises designed to verify your understanding of the instructions. After you completed the exercises, please raise your hand. Once everybody has completed the exercises, we will go over them together.

A.1 Instructions for Player A

You are participant A. Either you or participant B will decide how to divide 20 points between the four participants in your group. Being participant A, you can choose between four options. If you do not pass the decision to participant B, then participant B will not take a decision. You take the decision. If you delegate the decision to participant B, then you cannot take any further decision. Participant B will take the final decision. In the following table we show you again an overview over all distributions between which you (or, in case you delegate, participant B) can choose.

If you have chosen an allocation of the 20 dollars - or, in case you delegated, participant B has chosen an allocation of the 20 dollars, participant C and participant D are informed about

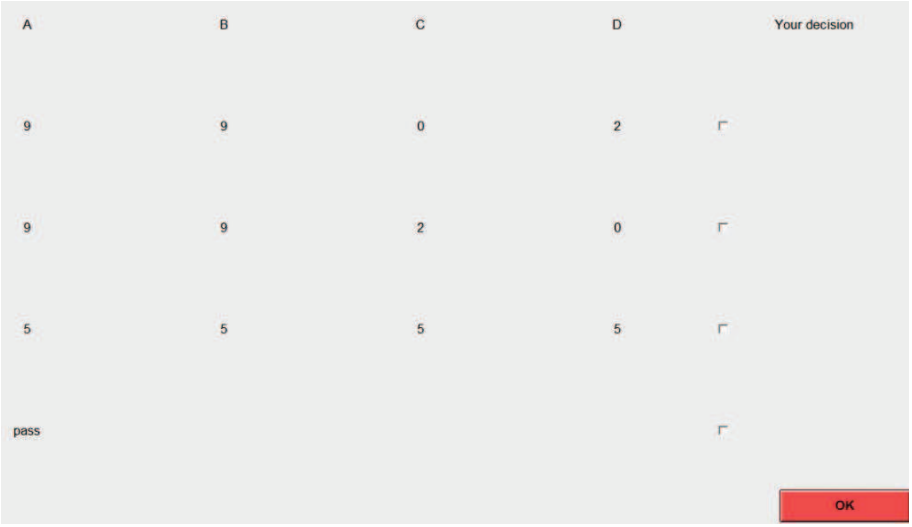
- whether you have delegated the decision or not, and
- what distribution was chosen.

		Your dollars	Dollars of B	Dollars of C	Dollars of D
You can choose	Distribution 1	9	9	0	2
	Distribution 2	9	9	2	0
	Distribution 3	5	5	5	5
	pass to B				
B can choose	Distribution 1	9	9	0	2
	Distribution 2	9	9	2	0

Then, either participant C or participant D is randomly chosen. This participant has the possibility to deduct dollars from you, participant B and the other participant (either D or C), at the cost of one dollar. The randomly chosen player can deduct a maximum of 7 dollars, but can also deduct less. The player can never deduct more dollars than the dollars you earned through the chosen allocation.

A.1.1 What will happen on the computer

Your introduce your decision on a screen as the following:



If you want to choose distribution 1, then you click the top small square on the right side. If you want to choose distribution 2, you click the second square. If you want to choose distribution 3, you click the third square. If you want to delegate the decision to participant B, you click in the last square.

After choosing one distribution, you click on the OK-button on the bottom right. As long as you don't click on this button, you can rethink your choice, and select something else. After you (and / or participant B) and the randomly chosen player C or D have made the decision, the experiment is finished and you get your final payoff paid in cash. To summarize, you only take one decision, which you have to introduce in the above screen. Think carefully about your decision. Do you have questions?

A.2 Instructions for Player B

You are participant B. Either participant A or you will decide how to divide 20 dollars between the four participants in your group. Participant A can choose between three distributions, or can pass the decision to you, participant B. If A passes the decision, you can only decide among two distributions. If A does not pass the decision, then you will not take a decision. If A delegates the decision to you, you will take the final decision. In the following table we show you again an overview over all distributions between which A (or you, in case A delegates) can choose.

- INPUT SAME TABLE AS FOR PLAYER A -

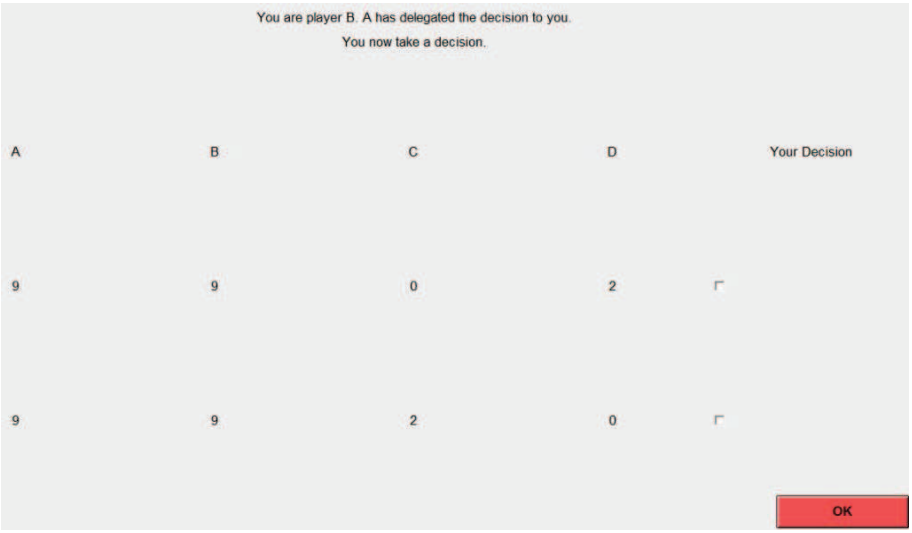
If A has chosen an allocation of the 20 dollars - or, in case A delegated, you have chosen an allocation of the 20 dollars, participant C and participant D are informed about

- whether A has delegated the decision or not, and
- what distribution was chosen.

Then, either participant C or participant D is randomly chosen. This participant has the possibility to deduct dollars from you, participant A and the other participant (either D or C), at the cost of one dollar. The randomly chosen player can deduct a maximum of 7 dollars, but can also deduct less. The player can never deduct more dollars than the dollars you earned through the chosen allocation.

A.2.1 What will happen on the computer

If participant A delegates the decision between distribution 1 and 2 to you, then you will see the following screen:



If you want to choose distribution 1, then you click the top small square on the right side. If you want to choose distribution 2, you click the second square.

After choosing one distribution, you click on the OK-button on the bottom right. As long as you don't click on this button, you can rethink your choice, and select something else. After you (and / or participant A) and the randomly chosen player C or D have made the decision, the experiment is finished and you get the final payoff paid in cash. To summarize, you only take one decision, which you have to introduce in the above screen. Think carefully about your decision. Do you have questions?

A.3 Instructions for Player R_1 / R_2

You are participant C. Either participant A or participant B will decide how to divide 20 dollars between the four participants. Participant A can choose between three distributions, or can pass the decision to participant B, who then can only decide among two distributions. If A does not pass the decision to participant B, then participant B will not take a decision. A takes the decision. If A delegates the decision to participant B, then participant B will take the final decision. In the following table we show you again an overview over all distributions between which A (or, in case A delegates, participant B) can choose.

- INPUT SAME TABLE AS FOR PLAYER A -

If A has chosen an allocation of the 20 dollars - or, in case A delegated, participant B has chosen an allocation of the 20 dollars, you or participant D are chosen randomly. The participant who is chosen has the possibility to deduct dollars from A, participant B and the other participant (either D or C), at the cost of one dollar. The randomly chosen player can deduct a maximum of 7 dollars, but can also deduct less. The player can never deduct more dollars than the dollars a participant earned through the chosen allocation.

A.3.1 Your Decision

Before you get to known which decision participant A and / or participant B has chosen and before you get to known if you or participant D were chosen to deduct dollars, we ask you to make a decision for each of the following five cases:

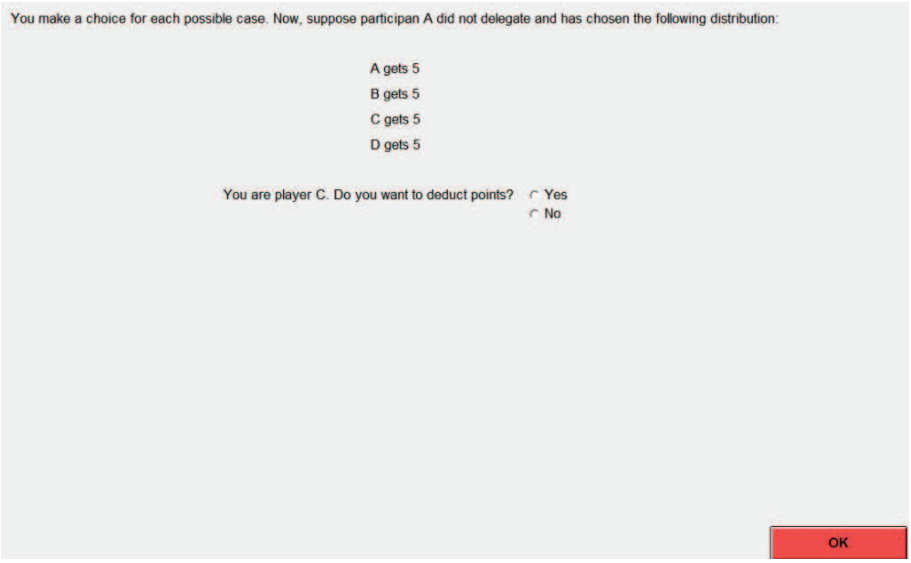
- Participant A does not delegate the decision and chooses allocation 1 (9,9,0,2)
- Participant A does not delegate the decision and chooses allocation 2 (9,9,2,0)
- Participant A does not delegate the decision and chooses allocation 3 (5,5,5,5)
- Participant A does delegate the decision and participant B chooses allocation 1 (9,9,0,2)
- Participant A does delegate the decision and participant B chooses allocation 2 (9,9,2,0)

In particular, this means that for each of the cases you have to say if you want to deduct dollars, and, if yes, how you want to distribute the deducted dollars on the other players. Participant A and / or participant B take the decision without knowing what you or participant D will do in each of the five cases.

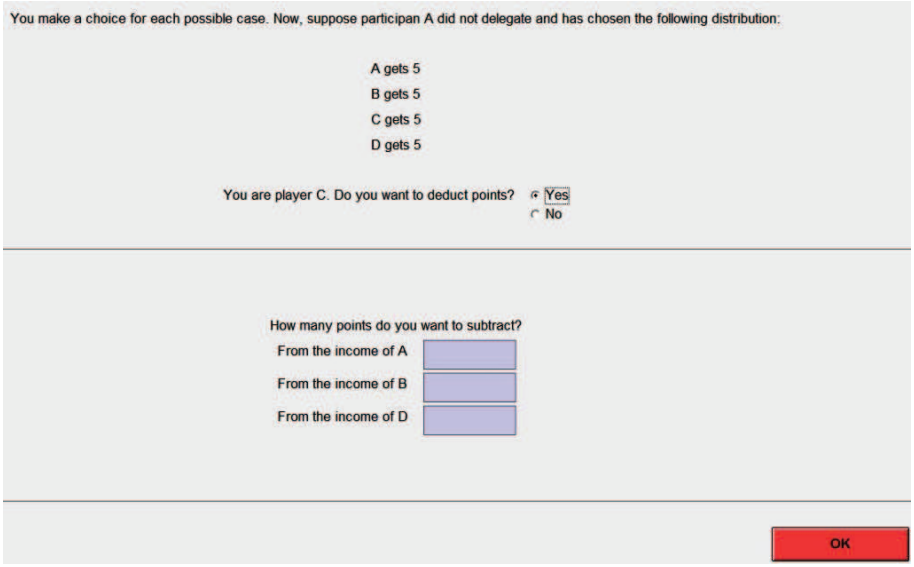
If you are chosen in the random selection process, then your decision is implemented for the case which results out of the decisions of participant A and / or participant B. Therefore, each of your five decisions can be determining for the final payments.

A.3.2 What will happen on the computer

Your decision in each of the cases you introduce in five screens like the following:

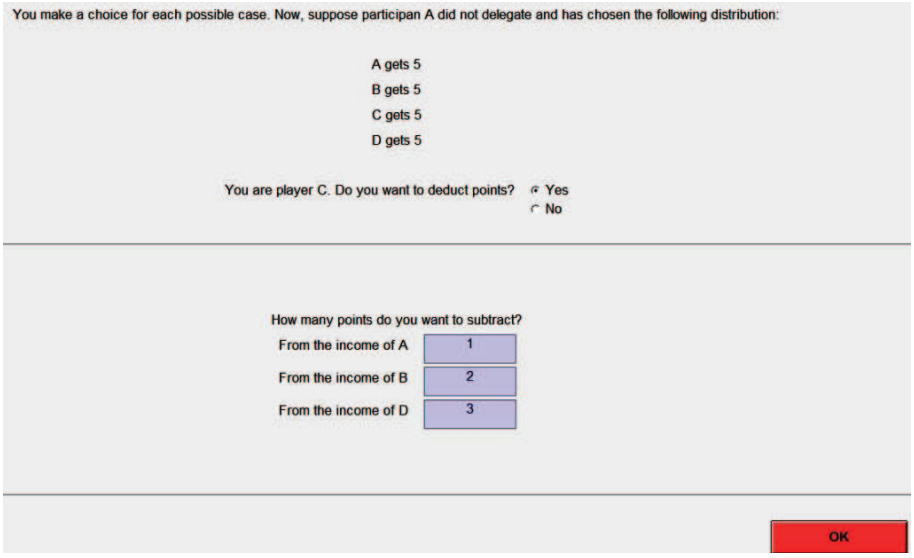


The above example-screen shows you the possibility "Participant A does not delegate the decision and chooses the allocation (5,5,5,5)". The screens for the other four cases look similar - please pay attention for which case you take the decision! If you click on "yes", then the following screen appears (see next page):



After clicking on "yes", you can introduce the respective amount you want to deduct in the three boxes . If you choose to deduct a dollar of at least one other player, you will be deducted one dollar - and the player of whom you want to deduct dollars loses the amount you stated in the respective field. If you choose "no" when asked if you want to deduct dollars, then the three small boxes do not appear (or disappear again), and you cannot deduct dollars.

After choosing one distribution, you click on the OK-button on the bottom right. As long as you don't click on this button, you can rethink your choice, and select something else.



In this example participant C wants to deduct dollars. Therefore, C clicked "yes", and the three small boxes appeared (see the picture on the previous page). Participant C deducts 1 dollar of participant A, 2 dollars of participant B, and 3 dollars of participant D. (This is just an example and not a suggestion of how you should act.) In total you can deduct up to 7 dollars in each case. You can as well (as in the above example) deduct less than seven dollars. Thereby, you cannot deduct more dollars of a participant than what that participant received according to the respective allocation. In the above example, therefore, you cannot deduct of any player more than 5 dollars.

When you press the OK-button, you come to the next case. As long as this button is not pressed, you can still change all your entries. Do you have any questions?

A.3.3 Examples

The examples and the exercises are the ones that were presented to player D. For I, R_1 and R_2, they were modified respectively.

- Example 1: Distribution 2 is chosen (either by yourself or by participant B) and the randomly chosen participant is participant C. C choses to give up one dollar, to deduct 3 dollars from you and 4 dollars from participant B. Then, the following distribution results:

	Your dollars	Dollars of B	Dollars of C	Dollars of D
Distribution 2	9-3 = 6	9-4 = 5	2 - 1 = 1	0

- Example 2: Distribution 3 is chosen (by yourself) and the randomly chosen participant is participant D. D choses to give up one dollar, to deduct two dollars from you, three dollars from participant B and one dollar of participant C. D choses to not to use the seventh possible dollar to deduct. Then, the following distribution results:

	Your dollars	Dollars of B	Dollars of C	Dollars of D
Distribution 3	5-2 = 3	5-3 = 2	5-1 = 4	5-1 = 4

- Example 3: Distribution 2 is chosen (either by yourself or by participant B) and the randomly chosen participant is participant D. D choses to not to give up one dollar, to deduct dollars from other players. Therefore, the resulting distribution is as explained above, (9,9,2,0).

- Example 4: Distribution 1 is chosen (by yourself or by participant B) and the randomly chosen participant is participant C. C choses to give up one dollar, to deduct two dollars from you, two dollars from participant B and two dollars of participant D. C choses to not to use the seventh possible dollar to deduct. Then, the following distribution results:

	Your dollars	Dollars of B	Dollars of C	Dollars of D
Distribution 1	9-2 = 7	9-2 = 7	0-1 = -1	2-2 = 0

A.3.4 Exercises Each Participant Had to Solve

Please answer the following questions. They only serve for helping you to get used to the experiment. The decisions and payments in the exercise here are chosen arbitrarily. Do not take them as suggestions for which allocation you should chose. Your answers here will not have any impact on the payments at the end of the experiment.

- Participant A has passed the decision to participant B. Whose decisions are relevant for the payments at the end of the experiment?

- Participant A did not pass the decision to participant B. Whose decisions are relevant for the payments at the end of the experiment?

- Distribution 2 is chosen. Participant C is randomly selected to deduct dollars, and wants to deduct the following bold printed amounts:

  Is this possible? If yes, please determine the resulting payment. If not, please make a mark where it is not possible.

	Your dollars	Dollars of B	Dollars of C	Dollars of D
Distribution 2	9	9	2	0
Deduction	**2**	**2**		**0**

- Distribution 2 is chosen. Participant D is randomly selected to deduct dollars, and wants to deduct the following bold printed amounts:

	Your dollars	Dollars of B	Dollars of C	Dollars of D
Distribution 2	9	9	2	0
Deduction	**2**	**1**	**3**	

  Is this possible? If yes, please determine the resulting payment. If not, please make a mark where it is not possible.

- Distribution 3 is chosen. Participant C is randomly selected to deduct dollars, and wants to deduct the following bold printed amounts:

	Your dollars	Dollars of B	Dollars of C	Dollars of D
Distribution 3	5	5	5	5
Deduction	**2**	**3**		**3**

  Is this possible? If yes, please determine the resulting payment. If not, please make a mark where it is not possible.

- Distribution 3 is chosen. Participant D is randomly selected to deduct dollars, and wants to deduct the following bold printed amounts:

	Your dollars	Dollars of B	Dollars of C	Dollars of D
Distribution 3	5	5	5	5
Deduction	**7**	**0**	**0**	

  Is this possible? If yes, please determine the resulting payment. If not, please make a mark where it is not possible.

- Distribution 1 is chosen. Participant D is randomly selected to deduct dollars, and wants to deduct the following bold printed amounts:

	Your dollars	Dollars of B	Dollars of C	Dollars of D
Distribution 1	9	9	0	2
Deduction	**0**	**0**	**0**	

  Is this possible? If yes, please determine the resulting payment. If not, please make a mark where it is not possible.

Please remember that these are only exercises, and that all numbers were chosen arbitrarily. You should not use the numbers to orient your decision on them. When you have solved the exercises, please raise your hand. You may think about your decision in the experiment.

A.3.5 Belief Elicitation

After the experiment had finished, in four of the fourteen sessions we elicited beliefs from the subjects with respect to the deduction behavior of the receivers. Subject A and B had to indicate for each of the five scenarios whether they expect player C to deduct points, and, if yes, how much they expect C to deduct from each of the respective subjects in his/her group. The screens used to elicit belief from the As were the following four:

Now that you have made your decision, we would like to give you a chance to earn some extra dollars.

We will ask you what you belief the choice of participant C will be and has been in previous sessions. If your answer is within +/- $1 of the true average punishment of the C's in previous sessions, you will earn an additional $5.

The average punishment is calculated using only the people that punished at least one person in the respective case. Hence, if you think player C did punish, we want you to state your beliefs about previous levels of punishment, conditional on punishing at all.

If you think that in a case C did not punish, this input counts as if you state C deducted 0 points.

If you have any questions, please raise your hand and wait for assistance.

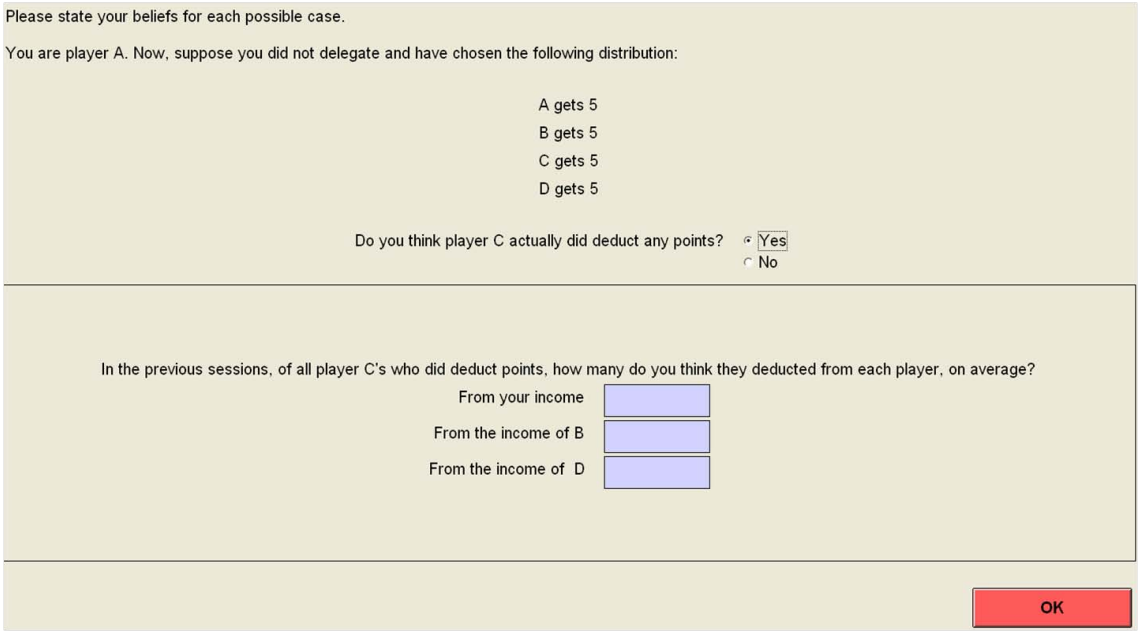
Please state your beliefs for each possible case.

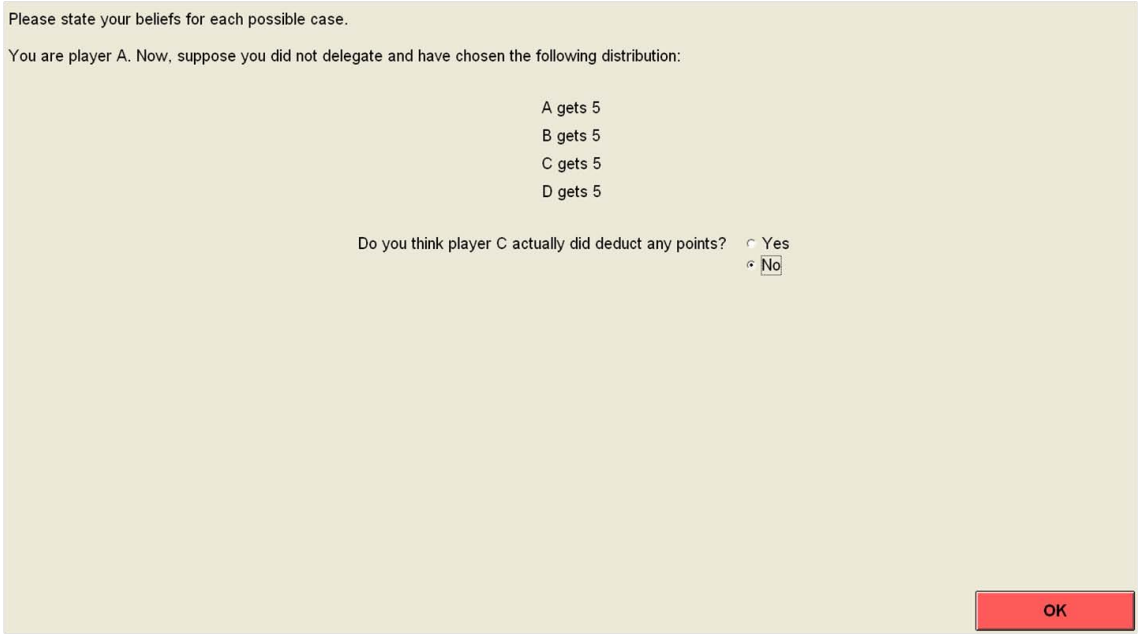
You are player A. Now, suppose you did not delegate and have chosen the following distribution:

A gets 5
B gets 5
C gets 5
D gets 5

Do you think player C actually did deduct any points? ○ Yes ○ No

OK





Instead of the latter three screens, player B saw personalized screens similar to the following one:

Please state your beliefs for each possible case.

You are player B. Now, suppose A delegated and you have chosen the following distribution:

A gets 9
B gets 9
C gets 2
D gets 0

Do you think player C actually did deduct any points? ○ Yes ○ No

OK