University of Heidelberg

Department of Economics

Discussion Paper Series | No. 627

Is fairness intuitive? An experiment accounting for the role of subjective utility differences under time pressure

Anna Merkel and Johannes Lohse

November 2016

Is fairness intuitive? An experiment accounting for the role of subjective utility differences under time pressure *

Anna Merkel[†] and Johannes Lohse[‡]

November 2016

Abstract

Economists are increasingly interested in the cognitive basis of pro-social behavior. Using response time data, several authors have claimed that "fairness is intuitive". In light of conflicting empirical evidence, we provide theoretical arguments showing under which circumstances an increase in "fair" behavior due to time pressure provides unambiguous evidence in favor of the "fairness is intuitive" hypothesis. Drawing on recent applications of the Drift Diffusion Model (Krajbich et al., 2015a), we demonstrate how the subjective difficulty of making a choice affects choices under time pressure and time delay, thereby making an unambiguous interpretation of time pressure effects contingent on the choice situation. To explore our theoretical considerations and to retest the "fairness is intuitive" hypothesis, we analyze choices in two-person prisoner's dilemma and binary dictator games. As in previous experiments, we exogenously manipulate response times by placing subjects under time pressure or forcing them to delay their decisions. In addition, we manipulate the subjective difficulty of choosing the fair relative to the selfish option across all choice situations. Our main finding is that time pressure does not increase the fraction of fair choices relative to time delay irrespective of the subjective difficulty of choosing the fair option. Hence, our results cast doubt on the hypothesis that "fairness is intuitive".

Keywords: distributional preferences, cooperation, response times, time pressure, cognitive processes, drift diffusion models

JEL Classification: C72, C91, C92, D03, D87

*We thank Christoph Vanberg, Gustav Tinghög, Ariel Rubinstein, Gert Pönitzsch, Tobias Pfrommer, as well as seminar participants at the University of Heidelberg, ZEW Mannheim, IfW Kiel, Thurgau Experimental Economics Meeting, IMEBESS Rome, Kings College London and ESA Bergen for their helpful comments.

†Corresponding author at: Department of Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany, anna.merkel@awi.uni-heidelberg.de

‡Department of Economics, University of Birmingham, J G Smith Building, Edgbaston Campus, J.Lohse@bham.ac.uk

1 Introduction

Economists are increasingly interested in understanding the cognitive (Alós-Ferrer and Strack, 2014) and emotional (Loewenstein, 2000; Hopfensitz and Reuben, 2009; Drouvelis and Grosskopf, 2016) processes that drive pro-social behavior. One of the central questions within this literature is whether "fair" behavior is intuitive and automatic or follows from a deliberative weighting of the costs and benefits of behaving fairly. Answering this question is empirically challenging, since the cognitive processes underlying overt behavior are not directly unobservable. A number of authors have approached this problem by analyzing response times. They argue that response time (i.e. the time it takes before a decision is made) is indicative of the amount of deliberation involved in the decision process (Rubinstein, 2007; Spiliopoulos and Ortmann, 2015). In particular, fast responses are seen to emerge from intuition, while slow responses are seen to involve more deliberative reasoning. To test the causal impact of deliberation on choices, a related approach attempts to exogenously manipulate the amount of deliberation. In particular, placing subjects under time pressure is seen as a method that exogenously increases their reliance on intuition relative to placing them under time delay (Wright, 1974; Rand et al., 2012).

Recently, these two methods have been applied to investigate the cognitive basis of "fair" behavior.[1] For example, several papers have found that shorter response times are associated with more cooperative behavior in public good games (Rand et al., 2012; Lotito et al., 2013; Nielsen et al., 2014) and more pro-social behavior in dictator games (Cappelen et al., 2014). Furthermore, Rand et al. (2012, 2014) find that average contributions in a public good game are higher when subjects are placed under time pressure as compared to subjects who are forced to delay their contribution decision.[2] Especially these latter results have inspired the "fairness is intuitive" (Cappelen et al., 2014) hypothesis. According to the "fairness is intuitive" (FII) hypothesis, a decision maker intuitively prefers the "fair" option, i.e. cooperation in a public good or the pro-social option in a dictator game. However, this predisposition towards fair behavior can be overridden by a more deliberative weighting of the costs and benefits of fair behavior, such that deliberation can lead to more "selfish" choices (Rand et al., 2012). A potential explanation for this finding - as stated in the social heuristics hypothesis (Rand et al., 2014) - could be that sharing resources with others is usually a successful strategy in repeated interactions outside the lab and therefore applied as a behavioral heuristic within the lab, even in one-shot games.

An important implication of the FII hypothesis is that choices made by the same decision maker may be fair or selfish, depending on the level of deliberation she engages in during the choice process. Providing robust evidence on the link between intuition and fair behavior might, thus, be important to better understand the variability of fair behavior at the individual level. Furthermore, such evidence may contribute towards a unified explanation for several experimental findings that are otherwise difficult to align with existing theories of pro-social behavior. These findings include the observation that people tend to be more inclined to give in the lab than in the field, the observation that a significant number of people tend to avoid being asked to give as well as the observation that delayed payments decrease giving (Dreber et al., 2014).

[1] Obviously, the economics literature has come up with various, sometimes competing, notions of what constitutes "fair" behavior in a given choice situation (Rabin, 1993; Engelmann and Strobel, 2004; Fehr and Schmidt, 2006). In order to highlight that the literature on deliberation and fairness draws on a variety of fairness concepts, we use quotation marks here. In section 2 we will describe in more detail what we refer to as a "fair" choice in the context of our paper.

[2] While subjects under time pressure and time delay contribute different amounts, they are found to hold similar beliefs regarding the average cooperation rates of other subjects. Hence, the observed changes in the fraction of "fair" choices cannot be explained by a change in beliefs.

In a recent meta-analysis, Rand (2016) finds that, on average, promoting intuition increases cooperative choices relative to promoting deliberation, as predicted by the FII hypothesis. However, this hypothesis has not been unequivocally confirmed in the empirical literature. For example, Piovesan and Wengström (2009) as well as Ubeda (2014), find that more pro-social choices in simple distribution tasks are associated with longer response times. Similarly, Lohse et al. (2016) find that making a contribution to a real-world public good is related to longer response times. Such inconclusive results can also be found in the literature which causally tests the FII hypothesis by placing subjects under time pressure during their choices. In contrast to the original results of Rand et al. (2012), Tinghög et al. (2013) and Verkoeijen and Bouwmeester (2014) do not find that time pressure increases the fraction of cooperative choices relative to time delay in one-shot public good games. Moreover, findings in Capraro and Cococcioni (2016) and Lohse (2016) suggest that placing subjects under time pressure may lead to more selfish choices in some instances. Taken together, these results could point towards a different behavioral model in which intuition promotes selfish behavior but where deliberation increases fairness concerns (e.g., Loewenstein and O Donoghue, 2007).

In light of this mixed evidence, we conduct a new test of the FII hypothesis. In this test we address a recent concern that factors other than intuition and deliberation also affect response times and thereby distort the identification of intuitive or deliberative choices from fast and slow responses. Examples of such factors are confusion (Recalde et al., 2014) as well as the subjective difficulty of making a choice (Evans et al., 2015; Krajbich et al., 2015a).[3] In this paper we focus on the latter factor and explore how the subjective difficulty of choosing between a fair and a selfish option affects choices under time pressure and time delay and thereby affects a causal test of the FII hypothesis.

Our proposed method builds on insights from a recent paper by Krajbich et al. (2015a) who use a so-called Drift Diffusion Model (DDM) to illustrate how differences in the perceived difficulty of a choice situation may affect individual response times. The DDM assumes that a decision maker evaluates her choice options according to an underlying value function the actual values of which she has to infer from stochastic signals. The central prediction of the DDM is that response times reflect choice difficulty, which is defined as a function of the decision maker's true underlying utility difference between the options of choice. If this utility difference is large, the decision maker will respond faster than if the underlying utility difference is small.

The DDM provides a plausible explanation for the observation that fair choices have been found to be faster or slower than selfish choices across different experiments: if, in a given experiment, the utility difference is larger for decision makers who prefer the fair option than for decision makers who prefer the selfish option, the DDM predicts that the fair option will be chosen faster than the selfish option. Hence, in such an experiment, fair choices will be associated with faster response times than selfish choices. On the other hand, if in a different experiment the incentives are such that the utility difference is smaller for decision makers who prefer the fair option than for decision makers who prefer the selfish option, fair choices will be slower than selfish choices. Krajbich

[3] One plausible link between confusion and response times could be that quick decisions are more prone to errors. For instance, Recalde et al. (2014) show that subjects who select a strictly dominated contribution level in a public good game with an interior equilibrium are faster than subjects selecting the dominant contribution level. Hence, the previous findings that fast choices are more cooperative in a linear public good game could also be due to confused subjects quickly selecting (dominated) positive contribution levels. While confusion could explain the negative correlation between response times and contributions found in some of the previous studies, it is unclear how confusion could account for the positive correlation between response times and contributions found in other studies which use the exact same games and do not alter the location of the equilibrium.

et al. (2015a) substantiate these theoretical arguments in a series of public good experiments. In games in which the benefits of cooperation (i.e. the MPCR) are sufficiently high, they replicate the finding that more cooperative choices (i.e. higher contributions) are correlated with shorter response times. While evidence of this kind has previously been interpreted as supporting the FII hypothesis (Rand et al., 2012), an interpretation based on the DDM would suggest that in this particular high benefit game, contributors find it easier to select their preferred option than free-riders. Consistent with this alternative explanation, Krajbich et al. (2015a) find that more cooperative choices are associated with longer response times, when the benefits of cooperation are low.[4] Taken together, this reversal of the correlation between response times and fair choices indicates that response times could indeed reflect the subjective difficulty of choosing the fair option instead of the intuitive and deliberative processes involved in the decision.

However, the observation that response times are affected by the perceived choice difficulty is not necessarily incompatible with the FII hypothesis because there is a crucial difference between the statement that a choice is "intuitive" and the statement that a choice is "easy". If multiple factors (Chen and Fischbacher, 2015) contribute to a decision maker's response time, the observation that fair choices are associated with longer response times may be driven by the subjective difficulty of making a choice in this particular situation. It might, however, still be the case that forcing subjects to rely on intuition in the same choice situation will make them more likely to choose the fair option, irrespective of the choice difficulty. This illustrates that, in theory, fair choices can be subjectively more difficult than selfish choices but intuition could still promote fair choices at the same time. Thus, the FII hypothesis could be entirely compatible with the finding that subjective utility differences also affect response times.[5] For example, the DDM and the FII could describe two separate decision processes that operate in parallel within a single decision maker and both influence the final decision being taken.[6]

While not directly contradicting the FII hypothesis, the findings of Krajbich et al. (2015a) still have major implications for a *causal* test of this hypothesis. Given that subjective utility differences have been shown to affect response times, they may *also* affect choices under time pressure and time delay. If true, the overall effect will depend on how time pressure affects choices according to the FII hypothesis *and* the DDM. According to the FII hypothesis, time pressure should always increase the fraction of fair choices. We use a standard version of the DDM to show that time pressure can decrease or increase the fraction of fair choices, depending on whether decision makers who prefer the fair option find it subjectively more or less difficult to make a choice compared to decision makers who prefer the selfish option. Hence, even without explicitly modeling the relative size of each effect or the exact timing, it is clear that there are choice situations in which both, the FII hypothesis and the DDM, predict that time pressure should increase the fraction of fair choices. If time pressure actually increases the fraction of fair choices in these situations, then this can only provide ambiguous evidence in favor of the FII hypothesis because the observed increase could be fully accounted for by the DDM. Instead, if no such increase is observed than this is unambiguous evidence against the FII hypothesis. At the same time, there are also situations in which the DDM and the FII hypothesis predict that time pressure affects the fraction of fair choices in opposite directions. In these situations, observing that time pressure increases the fraction of

[4] A similar reversal of correlations is also observed when the same authors analyze non-strategic decision tasks. In a series of binary allocation tasks, subjects who chose the pro-social option in a majority of trials are on average faster in trials in which they choose the pro-social option than in trials in which they choose the selfish option.

[5] In fact, Krajbich et al. (2015a) clearly state that their results simply speak against the use of response time correlations to identify the cognitive process but do not rule out the possibility that "fairness is intuitive".

[6] Alternative theories that also consider the possibility of multiple decision processes operating in parallel or sequentially are presented in Caplin and Martin (2015) and Alós-Ferrer (2016).

fair choices provides unambiguous evidence in favor of the FII hypothesis because the observed increase cannot be attributed to the DDM. Instead, if time pressure has no or even a negative effect, the FII hypothesis cannot be rejected because the positive effect of time pressure on the fraction of fair choices attributable to the FII hypothesis may be cancelled out by a negative effect attributable to the DDM. Hence, classifying whether the DDM and the FII hypothesis predict consistent or opposite effects of time pressure on the fraction of fair choices affects how to interpret the evidence. If it is not possible to classify a choice situation into one of those two types, the FII hypothesis could be spuriously accepted or rejected based on the observed effect of time pressure on choices. This last point is of central importance for understanding the mixed evidence in previous studies that have interpreted the effects of time pressure on fair choices without controlling for the effect of subjective utility differences.

In light of these considerations, our proposed experimental test comprises two elements: First, we compare treatments in which subjects are placed under time pressure with treatments in which subjects are forced to delay their decision, allowing us to *causally* test the FII hypothesis. Second, subjects take multiple choices across which we vary the subjective attractiveness of the fair option by increasing the social benefits of fair behavior. Specifically, our experiment includes choice situations in which we expect that decision makers who prefer the fair option will find it subjectively more, less or as difficult to choose as decision makers who prefer the selfish option. Thus, our experimental test is based on contrasting two different types of choice situations. In the first type, the DDM and the FII hypothesis both predict that time pressure increases the fraction of fair choices. In the second type, only the FII hypothesis makes this prediction, whereas the DDM predicts that time pressure should decrease the fraction of fair choices. To classify choice situations into one of the two types, we use an additional treatment, in which subjects are unconstrained in their response time and in which we observe response time correlations. According to Krajbich et al. (2015a), we should find that fair choices are correlated with shorter response times in decision problems in which fair choices are subjectively less difficult than selfish choices and vice versa.

In addition, our experiment comprises two further elements. First, it allows for a between- as well as a within-subject test of the FII hypothesis. Within-subject evidence is obtained by letting subjects take the same decision twice, once under time pressure and thereafter under time delay. This analysis serves as an additional robustness check by testing whether subjects indeed revise an initially fair choice and become more selfish when forced to delay their second decision. Second, in order to distinguish between fair choices in strategic and non-strategic decisions, we confront subjects in our experiment with a series of two-person prisoner's dilemma and binary dictator games. Thereby, we investigate whether pro-social behavior follows a common cognitive pattern across different contexts.

Overall, our test provides little empirical support for the hypothesis that "fairness is intuitive". In our between-subject comparison of choices, we find that even in games in which the correlational evidence would suggest that choosing the fair option is less difficult than choosing the selfish option, time pressure does not lead to a significant increase in fair choices as compared to time delay. Hence, even in situations in which the FII hypothesis and the DDM make the same prediction, we do not confirm that time pressure leads to a higher fraction of fair choices which is unambiguous evidence against the hypothesis that "fairness is intuitive". Similarly, we do not find evidence supporting the FII hypothesis in our within-subject analysis. Subjects who take their first decision under time pressure are as likely to revise an initially fair choice as subjects who take their first decision under time delay. Indeed, we observe a strong trend towards selfish behavior that is independent of the response time constraint as well as the benefits of choosing the fair option.

The remainder of the paper is organized as follows. Section 2 contains a detailed description of the DDM and summarizes our predictions. In section 3, we explain our experimental design. The results are summarized in Section 4. In section 5, we discuss our results and conclude.

2 Theory and Predictions

The FII hypothesis is based on a dual-process framework in which decisions are jointly determined by a fast and intuitive system I and a more deliberative and rather slow system II (Kahneman, 2003; Frederick, 2005). According to the "Social Heuristics Hypothesis" (Rand et al., 2014), the intuitive system I adopts strategies which are typically advantageous and successful in everyday life as automatic heuristics. The more deliberative system II is seen to be more sensitive to the actual details of a particular choice situation. Hence, deliberation can override generalized heuristics adopted by system I and cause decision makers to choose strategies which are advantageous in a specific context. Following this line of reasoning, Rand et al. (2014) argue that decision makers may develop a cooperation heuristic outside the lab where interactions are usually repeated and where cooperation can be an advantageous long-term strategy. Subjects may then apply this heuristic in the atypical context of one-shot laboratory experiments. Upon deliberation, however, they may realize that there are no strategic incentives to cooperate in this setting which leads to more defection. While the authors point to cooperation as one potential application of the "Social Heuristics Hypothesis", the same mechanism could apply more broadly to non-strategic decision problems such as choices in the dictator game. Under the assumption that outside the lab many interactions are repeated, sharing resources with other people is also advantageous because the same people may reciprocate generous behavior at a later point in time or because failing to share resources might be costly in terms of reputation. Hence, the "Social Heuristics Hypothesis" can be interpreted to predict that people are generally predisposed to behave fairly, i.e. to cooperate or to choose more pro-social outcomes. We summarize this claim that intuition promotes generally fair behavior as the FII hypothesis.

The FII hypothesis generates empirically testable predictions concerning the effect of time pressure and time delay on fair choices. Given that heuristics are seen to be relatively independent of the actual choice details, the FII hypothesis predicts that time pressure, by constraining deliberation, should always increase the fraction of fair choices relative to time delay.

However, as pointed out in section 1, observing a higher fraction of fair choices under time pressure cannot be unambiguously interpreted as evidence in favor of the FII hypothesis without accounting for the difficulty of the choice situation. To illustrate how the subjective difficulty of the choice situation could affect choices under time pressure and thereby confound a test of the FII hypothesis, we will describe the DDM in more detail.[7]

Assume that a decision maker faces a binary choice between a "fair" (henceforth: F) and a "selfish" (henceforth: S) option. The DDM assumes, that the decision maker is initially unaware of the utility value induced by each of these options, i.e. she does not know whether $u^i(S) > u^i(F)$, $u^i(S) < u^i(F)$ or $u^i(S) = u^i(F)$. However, the decision maker can accumulate stochastic information regarding her preferences until she is sufficiently sure which option she prefers. The evidence

[7] In particular, we will refer to versions of the DDM that have recently been applied to value-based choices and social dilemma situations (Polanía et al., 2014; Krajbich et al., 2014, 2015b). For a more extensive review of the behavioral foundations and the application of DDM in psychology refer to the descriptions in Ratcliff (1988), Ratcliff and Rouder (1998) and a recent summary of this topic aimed at economists by Clithero (2016).

accumulation process is assumed to take place in a series of time periods t. In each t, the decision maker observes two new stochastic value signals F_t and S_t which are independently drawn from two normal distribution functions with means $u^i(F)$ and $u^i(S)$. The difference between the two value signals observed in a period, $F_t - S_t$, is added to a subjective state variable X_t^i. The sign and size of X_t^i reflects the stochastic evidence accumulated until period t. Thus, X_t^i encodes the relative probability that F yields a higher utility than S based on the signals received up to period t (Krajbich et al., 2014; Caplin and Martin, 2015). The accumulation process continues until the subjective state variable crosses a pre-defined upper threshold $+a$, inducing the decision maker to choose F, or the lower threshold $-a$, inducing the decision maker to choose S. The assumption that a decision is made once either threshold is crossed can be interpreted as the decision maker being "confident enough" (Caplin and Martin, 2015) to make a choice. The length of the accumulation process, i.e. the number of time periods before reaching the upper or lower bound, corresponds to the decision maker's response time.[8]

The standard DDM we describe here generates two well-established predictions (e.g., Ratcliff and Rouder, 1998). First, the decision maker's response time crucially depends on the underlying utility difference, $|u^i(F) - u^i(S)|$. If this difference is large, the decision maker is expected to decide faster than if the underlying utility difference is small because, in expectation, she has to sample fewer signals for X_t^i to cross the threshold of her preferred option. Second, given that the final decision is reached by observing a series of noisy signals, a decision maker will be more likely to make a mistake (i.e. to choose the option that she does not prefer given her own preferences), the smaller the underlying utility difference between the fair and the selfish option. A small utility difference between the fair and the selfish option implies that the decision maker is more likely to receive signals that contradict her "true" preference. This, in turn, increases the likelihood of making a mistake by choosing the non-preferred option.

Jointly, these two properties of the DDM generate a third prediction concerning the effect of time pressure and time delay on choices. Time pressure forces individuals with otherwise longer response times to make a choice before being sure enough which option to choose. If we assume that this causes individuals to choose randomly, time pressure can increase the frequency of mistakes relative to time delay.[9] Importantly, the likelihood of making a mistake under time pressure is larger the smaller the subjective utility difference between the two options. Hence, under time pressure those decision makers whose underlying utility differences are smaller will be more likely to choose randomly compared to decision makers whose underlying utility differences are larger.

This third prediction has important implications for the effect of time pressure and time delay on the fraction of fair choices: in choice problems in which decision makers who prefer the fair option have larger utility differences than decision makers who prefer the selfish option, selfish decision makers are more likely to make a mistake than fair decision makers. Consequently, in these situations, time pressure increases the fraction of fair choices relative to time delay. The reason is that some actually selfish individuals will choose the fair option by mistake when forced to decide under time pressure. There is also the opposite case: if decision makers who prefer

[8] A more formal description of the DDM is contained in Appendix A, where we also derive the predictions of the subsequent paragraphs.

[9] For simplicity, we assume that decision makers who have reached neither decision threshold choose completely randomly between the options of choice. This reflects the view that the evidence accumulation process happens unconsciously and the decision maker is only informed about the final outcome so that she is completely unaware of her preferences until the decision thresholds are reached. In Appendix A we demonstrate that the predictions derived in this section are also valid under the weaker assumption that decision makers choose according to the evidence seen so far, i.e. at lower confidence.

Table 1: Testing the "Fairness is intuitive" hypothesis

		predicted effects of time pressure	
		FII $\uparrow$	FII $\uparrow$
		DDM $\uparrow$	DDM $\downarrow$
		(1)	(2)
observed effect of time pressure on the fraction of fair choices	$\uparrow$	ambiguous accept (1a)	unambiguous accept (2a)
	$\leftrightarrow$	reject (1b)	ambiguous accept (2b)
	$\downarrow$	reject (1c)	ambiguous accept (2c)

the fair option have smaller utility differences compared to decision makers who prefer the selfish option, then fair decision makers are more likely to make a mistake compared to selfish decision makers. In this second case, time pressure will decrease the fraction of fair choices relative to time delay because many actually fair decision makers will choose the selfish option by mistake when forced to decide under time pressure.

Assuming that responses under time pressure are affected by the relative use of intuition over deliberation *as well as* by the subjective difficulty of making a choice, these predictions have important implications for testing the FII hypothesis. Namely, in choice situations in which those decision makers who prefer the fair option have larger utility differences than those decision makers who prefer the selfish option, the DDM and the FII hypothesis both predict that time pressure should increase the fraction of fair choices relative to time delay. Therefore, if time pressure actually increases the fraction of fair choices in these situations, then this can only provide ambiguous evidence in favor of the FII hypothesis because the same observation could be fully accounted for by the DDM (see Table 1, 1a). Unambiguous evidence in favor of the FII hypothesis can only be obtained from situations in which an increased fraction of fair choices under time pressure can be clearly attributed to the FII hypothesis and not to the DDM. These are situations in which those decision makers who prefer the fair option perceive smaller utility differences and consequently have larger response times than those decision makers who prefer the selfish option. In these situations, time pressure should decrease the fraction of fair choices according to the DDM and increase the fraction of fair choices according to the FII hypothesis. Hence, if there is an increased fraction of fair choices under time pressure, then this observation can be clearly attributed to the FII hypothesis and, thus, constitutes unambiguous evidence (2a).

These considerations imply that there is only type of observation which would be inconsistent with the FII hypothesis. If time pressure has no or even a negative effect on the fraction of fair choices although the DDM and the FII hypothesis would both support this prediction, then this observation constitutes unambiguous evidence against the FII hypothesis (1b and 1c). However, observing no or even a negative effect is not necessarily inconsistent with the FII hypothesis in situations in which the DDM and the FII hypothesis predict that time pressure has opposing effects on the fraction of fair choices (2b and 2c). Here, the positive effect attributable to the FII hypothesis may be canceled out by a lager negative effect attributable to the DDM. Hence, the FII hypothesis cannot be rejected if there is no or even a negative effect of time pressure on the fraction of fair choices.

These theoretical arguments are of crucial relevance for the existing literature examining the *causal* effects of time pressure on choices. Importantly, if different cognitive processes jointly influence behavior an increased fraction of fair choices under time pressure does not necessarily constitute evidence in favor of the FII hypothesis. By the same line of reasoning, observing no or even a negative effect of time pressure does not necessarily constitute evidence against the FII hypothesis. Thus, without accounting for the effect of the subjective difficulty on choices under time pressure, the FII hypothesis may have been falsely accepted or falsely rejected in previous studies.

3 Experimental Design and Procedures

In this section we describe how the discussion in sections 1 and 2 informs our experimental treatments, parameters, and predictions. Thereafter, we briefly summarize the experimental procedures.

3.1 Experimental Design

In our laboratory experiment we confront subjects with a series of binary choice situations: four prisoner's dilemma (see Figure 1) and four binary dictator games (see Figure 2). To analyze the causal effect of deliberation on fair choices each decision is taken under an exogenous time constraint, i.e. either time pressure or time delay.

In each game, subjects are asked to choose between a "fair" and a "selfish" option. In line with the FII hypothesis (Rand et al., 2014), outlined in Section 2, we label a choice as "fair" if it implies sharing resources with another individual at own costs. Therefore, we label cooperation as the "fair" choice, whereas defection corresponds to the "selfish" choice in the prisoner's dilemma games. In the binary dictator games, subjects have to choose between an equal allocation, which yields the same payoffs to the dictator and the recipient, and an unequal allocation, which maximizes the dictator's payoff but reduces the recipient's payoff relative to the equal allocation. Therefore, the equal allocation corresponds to the "fair" and the unequal allocation corresponds to the "selfish" choice. Across the games, we increased the social benefits of choosing the fair option from VERY LOW to HIGH.[10]

If subjective utility differences reflect the benefits of choosing the fair option, we would expect that in the VERY LOW games decision makers who prefer the fair option should perceive smaller utility differences than decision makers who prefer the selfish option. In these games, the benefits of choosing the fair option are relatively small since the decision maker needs to sacrifice a high amount of her own payoff to increase the payoff of the other participant by only a small amount. Hence, these are games in which we expect that the DDM and the FII hypothesis predict opposite effects of time pressure, allowing for an unambiguous test of the FII hypothesis. By the same logic we expect that in the HIGH games, decision makers who prefer the fair option perceive larger utility differences than decision makers who prefer the selfish option. Here, decision makers need to give up only a small amount in order to increase the payoff of the other participant by a high amount. Hence, we expect that decision makers who prefer the fair option in these games will have larger utility differences than those decision makers who prefer the selfish option.

[10] For example, in the VERY LOW binary dictator game, choosing the fair (equal) option increases the recipient's payoff by 10 cents for every Euro that the dictator gives up relative to the selfish (unequal) option. In HIGH, the recipient receives 2.25 for every Euro that the dictator gives up relative to the selfish option.

Figure 1: Prisoner's dilemma games used in the experiment

VERY LOW

	Cooperate	Defect
Cooperate	3.10, 3.10	1, 5.10
Defect	5.10, 1	2, 2

LOW

	Cooperate	Defect
Cooperate	4, 4	1, 6
Defect	6, 1	2, 2

MEDIUM

	Cooperate	Defect
Cooperate	6, 6	1, 8
Defect	8, 1	2, 2

HIGH

	Cooperate	Defect
Cooperate	8, 8	1, 10
Defect	10, 1	2, 2

Figure 2: Binary dictator games used in the experiment

VERY LOW

Unequal	11, 0
Equal	1, 1

LOW

Unequal	9, 0
Equal	3, 3

MEDIUM

Unequal	10, 1
Equal	6, 6

HIGH

Unequal	15, 2
Equal	11, 11

Despite these theoretical considerations, it is hard to predict a priori if choosing the fair option will be perceived as subjectively more or less difficult than choosing the selfish option in a given game. To gain empirical insights into the subjective difficulty of choosing the fair and the selfish option in all eight experimental games we observe under time pressure and time delay, we collected additional observations for these games by conducting sessions in which subjects could decide without being constrained in their response times. Based on the previous finding that response times strongly reflect the relative difficulty of the choice situation (Krajbich et al., 2015a), we use these additional observations to classify games according to whether choosing the fair option is subjectively more difficult and hence slower or less difficult and hence faster than choosing the selfish option.

In order to elicit choices under time pressure and time delay, we used the following procedures in our experiment:

Part 1 of the experiment consisted of two successive blocks. In block 1, subjects were confronted with the four prisoner's dilemma games displayed in Figure 1 in randomized order. After each prisoner's dilemma game, subjects made choices in filler games (see Appendix B). Once subjects had completed block 1 and a short questionnaire, we confronted them with the exact same four prisoner's dilemma and filler games again in block 2. The games were presented in the same order in block 1 and 2 for each subject.[11]

Part 2 of the experiment also consisted of two successive blocks. In block 1, subjects were confronted

[11] We randomized the order in which the prisoner's dilemma games were displayed across sessions. The filler games were presented in the same order in all sessions.

Table 2: Experimental Design

		TIME PRESSURE	TIME DELAY	UNCONSTRAINED
PART 1	BLOCK 1 4 PDs + 4 Filler Games	≤ 12 seconds	> 12 seconds	no constraint
	BLOCK 2 4 PDs + 4 Filler Games	> 12 seconds	> 12 seconds	no constraint
PART 2	BLOCK 1 4 BDs + 12 Filler Games	≤ 6 seconds	> 6 seconds	no constraint
	BLOCK 2 4 BDs + 12 Filler Games	> 6 seconds	> 6 seconds	no constraint

This table summarizes the Experimental Design. Each cell displays the response time limit which subjects faced during their choice. We abbreviate prisoner's dilemma as "PD", and binary dictator game as "BD".

with the four binary dictator games displayed in Figure 2 in randomized order. Choices were elicited using the strategy vector method, i.e. both subjects in a pair made a choice before the computer randomly assigned them to the roles of dictator or recipient. After each binary dictator game, subjects took choices in three filler games (see Appendix B). Once subjects had completed block 1 and another short questionnaire, they were confronted with the same four binary dictator and filler games again in block 2. Again, the games were presented in the same order in both blocks.

To analyze the effect of time pressure on the fraction of fair choices, we randomly assigned subjects to one of three (between-subject) conditions, in which we implemented different response time constraints: In the Time Pressure condition, subjects were constrained to choose under time pressure in block 1 and forced to delay their decision in block 2. In the Time Delay condition, subjects were forced to delay their decision in both, block 1 and block 2. We also conducted an Unconstrained condition, in which subjects did not face an exogenous time limit in either block.

Comparing choices in block 1 between the Time Pressure and the Time Delay condition, allows us to test the between-subject predictions of the "fairness is intuitive" hypothesis, i.e. by comparing the choices of two groups of subjects who faced different response time constraints. By comparing how choices change between block 1 and 2 in the Time Pressure and Time Delay condition, we can conduct within-subjects tests, i.e. by analyzing the impact of response time constraints on the choices of the same subject. By investigating the response times of subjects in the Unconstrained condition, we test whether choosing the fair option was subjectively more, less or equally as difficult as choosing the selfish option in a given game. Our experimental design is summarized in Table 2.

Time pressure was set to 12 seconds for all prisoner's dilemma games in part 1 and was reduced to 6 seconds for all binary dictator games in part 2. These time limits correspond to the first quartile of the response time distribution in the Unconstrained condition.[12] Similarly, time delay was set to

[12] We used the first decision taken in the Unconstrained treatment to calculate response time quartiles. To our knowledge there is no common method according to which time pressure was defined in previous studies. For instance, subjects in Rand et al. (2012) were constrained to decide within 10 seconds which corresponds to the median response time in their response time correlation study. Buckert et al. (2014) define time pressure as 2/3 of the median response time in a Cournot game. Our analysis of response times in the Unconstrained treatment

12 seconds in part 1 and reduced to 6 seconds in part 2. To ensure compliance with our treatment, we forced subjects to delay their decision by displaying the choice buttons only after 12 seconds (6 seconds) had passed. Since compliance with time pressure cannot be enforced in the same way, we instead chose to incentivize compliance by informing subjects that they would lose their show-up fee of 3 Euro if they violated the time constraint in the decision chosen for payment.[13] The time (in seconds) was displayed as a countdown on each decision screen in both, the Time Delay and the Time Pressure condition.

The payoffs were displayed graphically as stacked and colored bars in all games (see Appendix B for further details). Thus, the payoffs associated with each option were easily accessible and comparable, even under time pressure.

At the end of part 1 we elicited the subjects' belief regarding the choices of other participants which allows us to test whether time pressure and time delay affected beliefs differently.[14] Subjects were paid an additional Euro for a correct guess. In addition, we elicited whether subjects were able to recall their initial choice in part 1 and 2. This was done in order to check whether there are heterogeneous effects among subjects who do and those who don't remember their initial choice, given that initial choices may create a strong anchor.

For each binary choice, subjects were randomly rematched in pairs and no feedback on their partner's choice was given until the very end of the experiment. One of the binary games was randomly drawn at the end of the experiment and subjects were paid according to their own and their partner's choice in this game.

3.2 Experimental Procedures

The experiment was conducted at the AWI Lab at the University of Heidelberg. The sessions for the Unconstrained condition were conducted in October 2015. The Time Pressure and the Time Delay conditions were conducted in parallel between December 2015 and January 2016. In total, 134 undergraduate and graduate students of all disciplines from the University of Heidelberg participated in our experiment (62 in Unconstrained, 34 in Time Delay and 38 in Time Pressure).[15] Subjects were recruited online via HROOT (Bock et al., 2014). We restricted our recruitment to subjects who had not participated in more than four experiments (and no experiment involving social dilemma or distribution tasks). The experiment was programmed in z-Tree (Fischbacher, 2007). Subjects received all instructions (reproduced in Appendix B) on the screen and questions were answered privately. Subjects could complete parts 1 and 2 at their own speed but all subjects started part 2 at the same time. We informed subjects about the choice of the other participant

revealed that the response time distribution of the 25% fastest decision makers was independent of the order in which the games were presented. Thus, the time limit in our study avoids heterogeneous effects across different order conditions. The response time distribution is displayed in the Appendix.

[13] One important limitation of previous studies has been that a large fraction of subjects violate the time constraints set by the experimenter which potentially reduces their explanatory power (Tinghög et al., 2013). In contrast to previous studies, we observed few violations of the time limit: Averaged over all decisions, the time pressure condition was violated in 0.5 percent of the decision tasks.

[14] In the Time Pressure treatment, subjects were constrained to indicate their belief within 12 seconds. In the Time Delay treatment, subjects could indicate their belief only after 12 seconds had passed.

[15] We performed a power analysis to calculate the required sample size. We chose the number of subjects such that we would be able to identify a moderate effect of time pressure on choices, i.e. up to 30% difference in the fraction of fair choices under time pressure relative to time delay, assuming $\alpha = 0.05$ and $1 - \beta = 0.8$. We believe that requiring a moderate effect size is appropriate for testing a theory that has been interpreted as a general statement regarding fundamental tendencies in human behavior.

in the game chosen for payment at the end of the experiment, but we did not reveal the matching or the choices in the other games. At the end of the experiment, all subjects were paid in private. The average earnings were 12 Euro, including a 3 Euro show-up fee.

4 Results

In this section, we report the results of our experiment. First, we will summarize the results of the Unconstrained condition. Hence, this section gives an overview of the response time correlations observed for each game. Thereafter, we report the results of the Time Pressure and Time Delay conditions which allow us to test the predictions of the "fairness is intuitive" hypothesis.

4.1 Unconstrained condition

The purpose of the Unconstrained condition was to identify situations in which the fair choice is faster (or slower) than the selfish choice and to calibrate the time constraint that was used in the subsequent Time Pressure and Time Delay conditions. As outlined in section 3, we expect that fair choices are associated with higher response times in the VERY LOW games and with lower response times in the HIGH games.

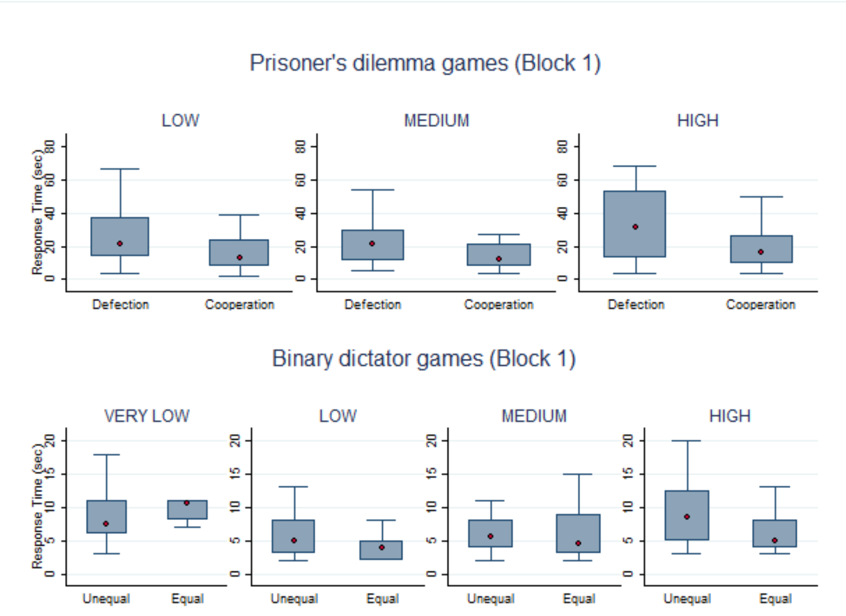
Our experiment included three prisoner's dilemma games for the Unconstrained sessions, namely the LOW, MEDIUM and HIGH game.[16] When confronted with these prisoner's dilemma games for the first time, we observe that the fraction of cooperators increases with the benefits: only 34 percent of subjects chose to cooperate in LOW, while 55 percent chose to cooperate in the MEDIUM and 58 percent in the HIGH game. The fraction of cooperative choices is significantly lower in the LOW game compared to the MEDIUM (Sign Test, $p < 0.05$) and HIGH ($p < 0.001$) games. The average rate of cooperation does not differ significantly between the MEDIUM and the HIGH game ($p = 0.72$). This observation confirms our prediction that the cooperative "fair" option is perceived as more attractive on average if the benefits of fair behavior are high.

In Figure 3, we compare the distribution of response times between cooperators and defectors for each of the three prisoner's dilemma games in block 1 (top panel). We find that the median response time of subjects who choose to cooperate is significantly shorter than the median response time of those subjects who choose to defect in each of the three games (Mann-Whitney U test, $p < 0.05$). When subjects are confronted with the same game for a second time in block 2, the median response time of cooperators is still lower than the median response time of defectors but this difference is more attenuated and only statistically significant in the LOW game (Mann-Whitney U test, $p < 0.01$). Contrary to the results in Krajbich et al. (2015a), we do not find that the correlation between response times and cooperative choices reverses as the benefits of cooperation increase from LOW to HIGH although cooperation rates increase from 34 to 58 percent. That is, even in the LOW game, we observe a negative and significant correlation between response times and the decision to cooperate (Spearmans Rho $= -0.29$, $p < 0.05$).

According to the DDM, these results suggest that for the chosen game parameters, our initial experiment only includes games in which cooperation is associated with larger utility differences

[16] We later added the VERY LOW game in the subsequent Time Pressure and Time Delay sessions for reasons we will discuss below.

Figure 3: Response times in prisoner's dilemma and binary dictator games (Unconstrained condition)



as compared to defection, i.e. fair decision makers find it subjectively less difficult to make a choice compared to selfish decision makers in all games. Given that in these cases the DDM and the FII hypothesis *both* predict that time pressure should increase the fraction of fair choices, the Unconstrained condition of our experiment only includes games from which we can obtain ambiguous evidence in favor of the FII (and unambiguous evidence against the FII). In order to also analyze the effect of time pressure in a game in which the DDM and the FII hypothesis are likely to make different predictions, we therefore added an additional prisoner's dilemma game (VERY LOW) in which we further reduced the benefits of cooperation to the virtually lowest possible level in the subsequent Time Pressure and Time Delay conditions.

In part 2 of the experiment, subjects were confronted with a series of four binary dictator games, across which we varied the benefits of giving to the other person from VERY LOW to HIGH.[17] As expected, the frequency of equal "fair" choices rises significantly from 9.7 percent in the VERY LOW game to 43.5 percent in the LOW, 51.6 percent in the MEDIUM, and to 61.3 percent in the HIGH game (Pairwise Sign Test, $p < 0.01$). Figure 3 (bottom panel) displays the relationship between fair choices and response times across all four games. In line with the results in Krajbich et al. (2015a), we observe that the correlation between choices of the fair allocation and response times depends on the benefits: the median response time of subjects who chose the equal option is larger than the median response time of subjects who chose the selfish option in the VERY LOW game (Mann-Whitney U test, $p < 0.1$). In the remaining three games, the median response times of those subjects who chose the equal option are smaller than the response times of those subjects who chose the selfish option. These response time differences are only weakly significant in the LOW and HIGH games (Mann-Whitney U test, $p < 0.1$) and insignificant in the MEDIUM game ($p = 0.64$). Hence, in contrast to the prisoner's dilemma games, the Unconstrained treatment

[17] In the Unconstrained condition the games were separated by additional distribution tasks, which were replaced by different filler games in the subsequent Time Pressure and Time Delay conditions. A full analysis of all 12 BD games is available upon request.

includes one binary dictator game (VERY LOW) in which, according to the DDM, the fair option is subjectively more difficult and, thus, associated with longer response times than the selfish option.

4.2 Constrained Response Time

This section summarizes the results from the Time Pressure (TP) and the Time Delay (TD) conditions. Our main goal is to test whether time pressure increases the fraction of fair choices between- and within-subjects, as predicted by the FII hypothesis. We use the results from the previous section to determine whether the DDM and the FII hypothesis predict time pressure effects that go in the same or in opposite directions. We refer to Table 1 from section 2 to interpret the evidence.

In a first step, we look at the between-subjects effect of time pressure in the prisoner's dilemma games by comparing initial, block 1, choices of subjects in the TP and the TD condition. Averaged across all four games, the frequency of cooperation does not differ between subjects in the TP (mean 46 percent) and the TD condition (mean 48 percent, Wilcoxon rank sum test, $p = 0.80$). Table 3 displays the effect of time pressure on the fraction of cooperative choices for each of the four games separately.

In section 4.1, we observed that in the LOW, MEDIUM and HIGH prisoner's dilemma games, fair choices are faster than selfish choices. Hence, these are situations in which the DDM and the FII hypothesis both predict that time pressure should increase the fraction of fair choices relative to time delay, i.e. the predicted effects go in the same direction. In contrast to Rand et al. (2012), we do not find that time pressure leads to a higher fraction of fair choices compared to time delay in any of these three games, as indicated by the p-Values reported in the last column of Table 3. We obtain the same result if we only analyze a subjects' first choice on the first decision screen (One-sided Fisher's exact test, LOW $p = 0.48$; MEDIUM $p = 0.14$; HIGH $p = 0.38$). This evidence stands in sharp contrast to the hypothesis that "fairness is intuitive". Given that both, the DDM and the FII, would predict that time pressure increases the fraction of fair choices, this is unambiguous evidence against the FII hypothesis.

We obtain the same results in the VERY LOW prisoner's dilemma game. Under time pressure, 39% of the subjects cooperate, under time delay this reduces to 38%. The difference is, however, not statistically significant, as indicated by the p-Value in Table 3. We obtain the same result if we analyze choices of subjects who took their very first choice in the VERY LOW game only (One-sided Fisher's exact test, $p = 0.68$). In this game, the observed absence of an effect may be due to the fact that the DDM and the FII hypothesis predict effects which affect fair choices in opposite directions. Thus, these opposite effects may cancel each other out, leading to no change in the overall observed fraction of fair choices.

One potential concern is that our time pressure manipulation could have affected beliefs. If subjects were more optimistic about average contributions of others in the Time Delay compared to the Time Pressure condition, we might have observed no evidence in favor of the FII hypothesis for this reason. To address this concern, we compare the stated beliefs in the Time Pressure and the Time Delay conditions. We find that in three out of four PD games, average beliefs did not differ between the two conditions. Only in the VERY LOW game, subjects in the Time Pressure condition believe that the average rate of cooperation is larger than subjects in the Time Delay condition (Mann-Whitney U test, $p < 0.05$). However, if at all, this difference in beliefs should

Table 3: Between-subject comparison of average cooperation rates
in prisoner's dilemma games

	Time Delay		Time Pressure		p-Value
	N	mean	N	mean	Fisher's exact (One-sided)
VERY LOW	34	0.38	38	0.39	0.55
LOW	34	0.47	38	0.39	0.34
MEDIUM	34	0.5	38	0.53	0.51
HIGH	34	0.59	38	0.55	0.47

The mean cooperation rate displayed is calculated over all orders. As our hypothesis suggests that the mean fraction of cooperative choices is higher in the time pressure than in the time delay condition, we report the p-Values of the one-sided Fisher's exact test.

make it more likely to find evidence in favor of the FII hypothesis.

In a second step, we analyze the within-subject effect of time pressure on choices in the prisoner's dilemma games. For this purpose, we compute switching probabilities by comparing a subject's first choice in block 1 with her second choice in the same game in block 2. If fairness was indeed intuitive, we would expect that more subjects initially choose to cooperate under time pressure and switch to defection under time delay. Whether the DDM predicts the same or the opposite switching pattern, depends on the subjective difficulty of choosing the fair or the selfish option: if subjects who prefer the fair option have larger underlying utility differences than subjects who prefer the selfish option, the DDM predicts that more subjects will switch from choosing the fair option under time pressure to choosing the selfish option under time delay than vice versa. This prediction is based on the same logic as our between-subjects predictions. In situations in which it is subjectively more difficult to choose the fair option, decision makers who prefer this option will be more likely to make a mistake under time pressure compared to decision makers who prefer the selfish option. When given enough time to deliberate, however, fair decision makers are more likely to revise their choice as compared to decision makers who prefer the selfish option. Hence, we would expect that more people switch from choosing the fair option under time pressure to choosing the selfish option under time delay. The opposite is true in situations in which those subjects who prefer the fair option have smaller utility differences than those subjects who prefer the selfish option. Here, we should find that more subjects switch from choosing the selfish option under time pressure to choosing the fair option under time delay.

Given that the likelihood to switch from one to the other option may also be due to the fact that subjects gain more experience with the task between their first and their second decision, we compare the switching rates within the Time Pressure condition to the switching rates in the Time Delay condition, where subjects take both decisions under time delay. To analyze switching behavior across the Time Delay and the Time Pressure condition, we computed a variable that takes a value of 1 if a subject switched from cooperation in block 1 to defection in block 2, and a value of -1 if a subject switched from defection in block 1 to cooperation in block 2 (see Table 4).

In section 4.1, we observed that in the LOW, MEDIUM and HIGH prisoner's dilemma games, fair choices are faster than selfish choices. Hence, these are games in which both, the DDM as well as the FII hypothesis predict that more subjects should switch from choosing the fair option under time pressure to choosing the selfish option under time delay than vice versa. The first thing to note is that within the Time Pressure condition, we do not find evidence that subjects are indeed more likely to switch from cooperation to defection in the LOW and the HIGH game. Only in the MEDIUM game, we find that significantly more subjects indeed cooperate initially and then switch to defection. The results of the Wilcoxon Sign Rank test are reported in Table 4. However, when we compare the switching behavior across the Time Pressure and the Time Delay condition, we do not find that subjects in the Time Pressure condition were more likely to switch from cooperation to defection as compared to decision makers in the Time Delay condition in any of the three games. Hence, instead of reflecting a reassessment of an initial intuitive decision, the decline of cooperative choices in the Time Pressure condition might simply reflect the well-known fact that subjects tend to become more selfish in repeated decisions even without receiving feedback (Ledyard et al., 1997). Therefore, our within-subject evidence in these three games does not support the FII hypothesis.

We obtain the same result in the VERY LOW game. Within the Time Pressure condition, subjects are indeed more likely to switch from choosing the fair option under time pressure to choosing the selfish option under time delay, as predicted by the "fairness is intuitive" hypothesis. However, when we compare the switching behavior across the Time Pressure and the Time Delay condition, we do not find that subjects in the Time Pressure condition are more likely to switch in the predicted direction compared to subjects in the Time Delay condition.

Taken together, we do not find evidence in favor of the hypothesis "fairness is intuitive", neither between- nor within-subjects. Moreover, finding no increase in fair behavior in decision problems in which the correlational evidence suggest that it is subjectively easy to choose the fair option, allows us to unambiguously reject the FII hypothesis.

Result 1: *In the Prisoner's dilemma games, we do not find that time pressure leads to a higher fraction of cooperative choices in comparison to time delay when comparing choices between-subjects. Within-subjects, we find that subjects in the Time Pressure treatment are as likely to revise an initially fair choice as subjects in the Time Delay condition.*

We now turn to analyzing the effect of time pressure on the fraction of fair choices in the binary dictator games. Averaged over all decision problems, we do not find evidence that subjects in the time pressure condition chose the equal "fair" option significantly more often than subjects who took their decision under time delay (43 vs. 50 percent, Mann-Whitney U test, $p = 0.37$). The

Table 4: Within-subject comparison of switching behavior in the prisoner's dilemma games

	Time Delay			Time Pressure			p-Value
	N	mean	Signrank p-Value	N	mean	Sign rank p-Value	Rank sum test
			(within-subjects)			*(within-subjects)*	*(across conditions)*
VERY LOW	34	0.26	0.00	38	0.24	0.03	1
LOW	34	0.12	0.28	38	0.08	0.37	0.74
MEDIUM	34	0.09	0.41	38	0.29	0.01	0.17
HIGH	34	0.21	0.05	38	0.13	0.12	0.62

In this table we report the direction of switches for each of the four prisoner's dilemma games. The switching variable takes a value of 1 if the subject switched from cooperation in block 1 to defection in block 2. The decision in block 1 is taken under time pressure in the TP condition and under time delay in the TD condition. The decision in block 2 is taken under time delay in both the TD and the TP condition. Columns 4 and 7 report the p-Value of a Wilcoxon Sign Rank Test, performed on subject's switching behavior within one condition. The last column reports the p-Value of a Wilcoxon rank sum test, which compares switching behavior across conditions.

Table 5: Between-subject comparison of the average rate of fair choices in the binary dictator games

	Time Delay		Time Pressure		p-Value
	N	mean	N	mean	Fisher's exact (One-sided)
VERY LOW	34	0.18	38	0.11	0.30
LOW	34	0.47	38	0.37	0.55
MEDIUM	34	0.62	38	0.61	0.26
HIGH	34	0.74	38	0.66	0.32

The mean rate of fair choices displayed is calculated over all orders. As our hypothesis suggests that the mean fraction of fair choices is higher in the time pressure than in the time delay condition, we report the p-Values of the one-sided Fisher's exact test.

results of our between-subject comparison are summarized in Table 5 in which we report the mean fraction of equal choices in each of the four binary dictator games separately.

Based on the correlational evidence, reported in section 4.1, we expect that in the LOW and the HIGH game, the DDM and the FII hypothesis both predict that time pressure should increase the fraction of fair choices relative to time delay. In none of these binary dictator games, we find that the equal "fair" allocation is chosen significantly more often in the time pressure condition than in the Time Delay condition. We obtain the same results if we use a subject's first decision only (One-sided Fisher's exact test; LOW $p = 0.18$, HIGH $p = 0.44$). These results do not support the hypothesis that "fairness is intuitive". Indeed, given that both, the FII hypothesis as well as the DDM would support the prediction that time pressure increases the fraction of fair choices, observing no effect constitutes unambiguous evidence against the FII hypothesis. However, observing no effect can only constitute ambiguous evidence in favor of the FII hypothesis.

For the VERY LOW game, our correlational evidence suggested that fair choices are associated with longer response times compared to selfish choices. Hence, this is a game in which the DDM and the FII hypothesis predict that time pressure affects choices in opposing directions. In this game, we find that the average rate of fair choices is 11% under time pressure and 18% under time delay. As indicated by the p-Value reported in Table 5, this difference is not statistically significant. We obtain the same result if we analyze the choices of those subjects who took their very first decision in the VERY LOW condition only (One-sided Fisher's exact test, VLOW $p = 0.29$). Given that for the VERY LOW game, the DDM and the FII hypothesis predict opposite effects of time pressure and possibly cancel each other out, we cannot unambiguously reject the FII hypothesis based on this observation.

As for the analysis of the prisoner's dilemma games, we test the FII hypothesis within-subjects by comparing switching behavior across the Time Pressure and the Time Delay conditions (see Table 6). For our interpretation, we distinguish between games in which the DDM and the FII hypothesis predict that time pressure affects choices in the same (LOW and HIGH) or opposite directions (VERY LOW game).

First, we do find that subjects are indeed more likely to switch from choosing the fair option under time pressure to choosing the selfish option under time delay than vice versa in the HIGH but not in the LOW game. The latter finding is inconsistent with the FII hypothesis. To control for a potential time trend in the probability to behave fairly which is not explained by our treatment, we compare the switching rate in the HIGH game across the Time Pressure and the Time Delay condition. The results of the Mann-Whitney U test are reported in Table 6. Our analysis shows that in the HIGH game, subjects in the Time Pressure condition were indeed more likely to switch from the equal to the unequal option compared to subjects in the Time Delay condition at a significance level of 10 percent. However, given that in the HIGH game the DDM and the FII hypothesis both make this prediction, this is at best ambiguous evidence in favor of the FII hypothesis.

In the VERY LOW game, we do not find that subjects are more likely to switch from choosing the fair option under time pressure to choosing the selfish option under time delay. However, given that in this game the DDM and the FII hypothesis predict opposing switching patterns, this finding is not necessarily inconsistent with the FII hypothesis. Hence, based on the evidence in the VERY LOW game, we cannot reject that FII.

Result 2: *In the binary dictator games, we do not find that time pressure significantly increases the fraction of fair choices relative to time delay in any of the games. Within-subjects, we find that subjects in the Time Pressure condition are more likely to switch from choosing the fair to choosing the selfish option compared to subjects in the Time Delay condition in one of four games. For this particular game this observation provides only ambiguous evidence in favor of the FII hypothesis.*

A potential concern, given the observed decline in average fair behavior across decisions, is that subjects' choices as well as their response time may be influenced by the order in which the different games were presented. This concern should already be limited by the fact that we presented the games in a randomized order. In addition, we did not find any evidence in favor of the hypothesis that "fairness is intuitive" using only the first decision taken by each subject. However, given that we have few initial choices for each order condition, we additionally address this concern using a set of probit regression models. Each of these models takes individual choices in the four decisions of block 1 as a dependent variable. Importantly, the dependent variable encodes the order in which the choices were taken. That is, if a subject entered the LOW game on the first screen, it is coded as choice 1 and if the LOW game appeared on the third screen, it is coded as choice 3. The results for the prisoner's dilemma games are reported in Table 7. In specification (1), we find no evidence

Table 6: Within-subject comparison of switching behavior in the binary dictator games

	Time Delay			Time Pressure			p-Value
	N	mean	Signrank p-Value *(within-subjects)*	N	mean	Signrank p-Value *(within-subjects)*	rank sum test *(across conditions)*
VERY LOW	34	0.26	0.32	38	0.08	0.18	0.81
LOW	34	0.12	0.13	38	0.11	0.25	0.75
MEDIUM	34	0.09	0.71	38	0.05	0.56	0.83
HIGH	34	0	1	38	0.16	0.03	0.10

In this table we report the direction of switches for each of the four binary dictator games. The switching variable takes a value of 1 if the subject switched from choosing the equal option in block 1 to choosing the unequal option in block 2. The decision in block 1 is taken under time pressure in the TP condition and under time delay in the TD condition. The decision in block 2 is taken under time delay in both the TD and the TP condition. Columns 4 and 7 report the p-Value of a Wilcoxon Sign Rank Test, performed on subject's switching behavior within one condition. The last column reports the p-Value of a Wilcoxon rank sum test, which compares switching behavior across conditions.

that time pressure increases the frequency of cooperation, when controlling for order effects and differences in the benefits of cooperation. Specification (2) contains additional interaction terms between the treatment dummy and the benefits associated with cooperation. Again, we find no evidence that time pressure significantly affects cooperative choices. Moreover, all interaction terms are insignificant. This is further evidence that even in those games where the cooperative choices is classified as easy (LOW, MEDIUM and HIGH), time pressure does not increase the frequency of cooperation. Finally, in specification (3) we add interaction terms between the TP and the SCREEN variables. These interactions terms would be significant if the order in which the games were presented would moderate the treatment effect - which we do not observe. Hence our results regarding a potential effect of time pressure are not driven by order effects. We do observe in all 3 specifications that the coefficients for the screen variables are significant. Thus, if a game was presented on a later decision screen, the likelihood that a subject would cooperate decreases relative to a game that was presented on the first decision screen.

Table 8 contains the results of a probit regression for choices in the binary dictator games. We use the same model and specifications as for the prisoner's dilemma games. We find only weak evidence for the presence of order effects. Importantly, in none of the three specifications there is a significant effect of time pressure on the likelihood of choosing the fair option. Moreover, the SCREEN variables, which capture potential order effects, are insignificant in most specifications and only one out of three interaction terms with the time pressure variable is significant. All in all, our robustness checks give little indication that Result 1 and 2 could be strongly influenced by order effects.

5 Conclusion and Discussion

In this paper, we propose and conduct a new test of the "fairness is intuitive" (Rand et al., 2012; Cappelen et al., 2014) hypothesis. Our test takes into account that a causal test of this hypothesis, using time pressure and time delay manipulations, also needs to account for the subjective difficulty of making a choice. We use a simple version of the Drift Diffusion Model (DDM) to show that time pressure can increase or decrease the fraction of fair choices, depending on whether decision makers who prefer the fair option perceive smaller or larger utility differences than decision makers who prefer the selfish option. Hence, these predicted effects may either be aligned with those of the FII hypothesis or affect choices under time pressure in the opposite direction. In our experiment, we then analyze the effect of time pressure in choice situations in which both, the DDM and the FII hypothesis, predict that time pressure should increase the fraction of fair choices. In neither of the relevant prisoner's dilemma or binary dictator games, we find that time pressure increases the fraction of fair choices which is unambiguous evidence against the FII hypothesis. Our experiment also includes choice situations in which we the DDM and the FII hypothesis predict opposite effects of time pressure, which may even cancel each other out. In the respective prisoner's dilemma and binary dictator games, we do not find evidence that time pressure has an effect on the fraction of fair choices. Although we cannot reject the FII hypothesis based on this observation, the observed evidence does certainly not provide clear evidence in favor of the FII hypothesis. Thus, overall our empirical test provides little support for the hypothesis that "fairness is intuitive". Importantly, this result holds between- and within-subjects.

On the one hand our rejection of the FII hypothesis is in line with a number of recent papers which also suggest that in some instances behaving fairly might not be intuitive and might even

Table 7: Random effects probit regression for the effect of time pressure in the prisoner's dilemma games

	(1)	(2)	(3)
TIME PRESSURE	-0.0727	0.0663	0.391
	(-0.23)	(0.15)	(0.71)
LOW (1=YES)	0.181	0.373	0.639
	(0.60)	(0.92)	(1.44)
MEDIUM (1=YES)	0.519*	0.481	0.803*
	(1.66)	(1.18)	(1.77)
HIGH (1=YES)	0.765***	0.916**	0.866**
	(2.88)	(2.37)	(2.25)
SCREEN2	-0.955***	-0.970****	-1.309***
	(-3.27)	(-3.29)	(-3.07)
SCREEN3	-0.796***	-0.818***	-0.625*
	(-3.08)	(-3.13)	(-1.66)
SCREEN4	-1.007****	-1.022****	-0.721*
	(-3.41)	(-3.44)	(-1.69)
TIME PRESSURE * LOW		-0.362	-0.876
		(-0.71)	(-1.44)
TIME PRESSURE * MEDIUM		0.0853	-0.518
		(0.17)	(-0.82)
TIME PRESSURE * HIGH		-0.276	-0.200
		(-0.54)	(-0.38)
TIME PRESSURE * SCREEN2			0.657
			(1.13)
TIME PRESSURE * SCREEN3			-0.348
			(-0.68)
TIME PRESSURE * SCREEN4			-0.556
			(-0.96)
CONSTANT	0.275	0.213	0.0343
	(0.86)	(0.59)	(0.09)
Observations	288	288	288
Subjects	72	72	72
Prob > Chi^2	0.0013	0.0067	0.0101

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

The dependent variable takes a value of 1 if cooperation was chosen and 0 otherwise. The variable Time Pressure equals 1 if the subject was assigned to the Time Pressure condition and 0 otherwise. Low, Medium and High are dummy variables for the games with which we confronted subjects in block 1. The coefficients report effects relative to the Very Low game which is the omitted category. The screen variables capture potential order effects by indicating whether a choice was presented on the second, third or last screen. The coefficients report effects relative to the decision on the first screen which is the omitted category.

Table 8: Random effects probit regression for the effect of time pressure in the binary dictator games

	(1)	(2)	(3)
TIME PRESSURE	-0.324	-0.729	0.0335
	(-0.89)	(-1.19)	(0.04)
LOW	1.656****	1.545***	1.787***
	(4.23)	(3.20)	(3.16)
MEDIUM	2.402****	2.086****	2.332****
	(5.82)	(4.21)	(4.09)
HIGH	2.555****	2.391****	2.510****
	(6.64)	(5.03)	(5.01)
SCREEN2	-0.444	-0.451	-0.211
	(-1.52)	(-1.53)	(-0.50)
SCREEN3	0.164	0.156	0.783*
	(0.60)	(0.56)	(1.84)
SCREEN4	0.0243	-0.00955	0.533
	(0.07)	(-0.03)	(1.02)
TIME PRESSURE * LOW		0.296	-0.0212
		(0.46)	(-0.03)
TIME PRESSURE * MEDIUM		0.710	0.461
		(1.06)	(0.59)
TIME PRESSURE * HIGH		0.457	0.433
		(0.67)	(0.61)
TIME PRESSURE * SCREEN2			-0.465
			(-0.78)
TIME PRESSURE * SCREEN3			-1.153**
			(-2.00)
TIME PRESSURE * SCREEN4			-1.022
			(-1.35)
CONSTANT	-1.627****	-1.462***	-1.961****
	(-3.82)	(-3.14)	(-3.46)
Observations	288	288	288
Subjects	72	72	72
Prob > Chi^2	0.0000	0.0000	0.0000

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

The dependent variable takes a value of 1 if the equal option was chosen and 0 otherwise. The variable Time Pressure equals 1 if the subject was assigned to the Time Pressure condition and 0 if the subject was assigned to the Time Delay condition. Low, Medium and High are dummy variables for the games with which we confronted subjects in block 1. The coefficients report effects relative to the Very Low game which is the omitted category. The screen variables capture potential order effects by indicating whether a choice was presented on the second, third or last screen. The coefficients report effects relative to the decision on the first screen which is the omitted category.

require additional deliberation or stronger self-control (Kocher et al., 2012; Fiedler et al., 2013; Tinghög et al., 2013; Martinsson et al., 2014; Duffy and Smith, 2014; Verkoeijen and Bouwmeester, 2014; Achtziger et al., 2015; Lohse, 2016; Capraro and Cococcioni, 2016; Tinghög et al., 2016). On the other hand, our results are surprising at least to the degree that they contradict a significant number of previous studies which tend to find that time pressure or other forms of inducing intuitive decision making lead to more cooperative or fair choices. For instance, a recent meta-study finds that relying on intuition relative to deliberation increases the average rate of cooperation by 6.1 percentage points in one-shot games (Rand et al., 2016a). Similarly, several current theories on the link between intuition and pro-social behavior are based on the idea that deliberation can never increase cooperation (Dreber et al., 2014; Rand et al., 2014; Bear and Rand, 2016)[18]. The observation that some experiments have found an increase of fair behavior under time pressure while other experiments report neutral or even negative effects, could well be in line with our theoretical considerations because none of the previous experiments has explicitly accounted for subjective utility differences. Hence, it is possible that some experiments have looked at choice situations where the DDM and the FII predict effects of time pressure which go in the same direction while other experiments have looked at choice situations in which the DDM and the FII hypothesis make opposite predictions. The most obvious reason for such differences is the choice of the experimental task or its parameters. But even in experiments that look at the same game (e.g., a public good game with MPCR 0.5) subject pools might differ (e.g., students vs. non-students) and it is conceivable that subjects with different individual attributes or cultural backgrounds might attach different subjective valuations to options in the same task.

Our theoretical considerations also shed some light on the discussion whether subjects that violate the time constraint should be excluded from analysis, as it has been done in prior time pressure experiments (Rand et al., 2012, 2016a). While this practice has already been criticized for violating random assignment (Tinghög et al., 2013)[19], our theoretical arguments show that resulting selection effects might depend on the choice situation. If decision makers perceive a given choice situation as difficult (i.e. they have smaller subjective utility differences) they are more likely to violate the time constraint. Hence, excluding these participants will lead to selection effects, the direction of which reflects whether "fair" decision makers find the choice more difficult than "selfish" subjects or vice versa.

At this point it is also important to stress that our paper does not attempt to directly replicate previous test of the FII hypothesis or pinpoint any other moderating factor (e.g. confusion, experience, social value orientation) that might also affect the direction of a time pressure effect. Rather, we aim at pointing out that it is unclear whether previous tests provide ambiguous or unambiguous evidence in favor of or against the FII hypothesis, since they do not account for subjective utility differences. Therefore our test differs from these previous tests of the same hypothesis along several dimensions that are motivated by our theoretical considerations: in our test subjects were confronted with several one-shot choice situations instead of only one, the specifics of each choice situation were only revealed on the decision screen and not on a preceding instruction screen[20], stakes were considerably higher than in many of the previous internet experiments, we used a graphical interface to visualize the payoffs of the different choice options and the compliance with the response time manipulations was more strongly enforced. We believe that each of these

[18] For a discussion of the last paper see Myrseth and Wollbrant (2016).

[19] For further issues with inferring compliance from absolute response times see the discussion in Myrseth and Wollbrant (2015).

[20] This is in line with Fiedler et al. (2013) and Capraro and Cococcioni (2016) but differs from (Rand et al., 2012). We, however, believe that giving subjects a possibility to fully deliberate on a task before entering the decision screen will affect the chances of isolating intuitive tendencies via time pressure.

design changes was well motivated and necessary in order to provide an unbiased test of the FII hypothesis. Furthermore, none of these changes should make it less likely to find evidence in favor of the FII hypothesis if it was generally valid as suggested by the mechanisms behind the social heuristics hypothesis (Rand et al., 2014).

Overall our results suggest that the link between intuition and fairness is more complicated and nuanced than previously thought. A closer inspection of other moderating factors might provide useful insights into the conditions or individual attributes that influence the link between intuition and fairness. First steps into this direction have already provided first insights into the role of confusion (Recalde et al., 2014; Stromland et al., 2016), gender (Rand et al., 2016b; Tinghög et al., 2016), culture (Nishi et al., 2016) and social-value-orientation (Chen and Fischbacher, 2015; Mischkowski and Glöckner, 2016).

References

A. Achtziger, C. Alós-Ferrer, and A. K. Wagner. Money, depletion, and prosociality in the dictator game. *Journal of Neuroscience, Psychology, and Economics*, 8(1):1, 2015.

C. Alós-Ferrer. A dual-process diffusion model. *Journal of Behavioral Decision Making*, 2016.

C. Alós-Ferrer and F. Strack. From dual processes to multiple selves: Implications for economic behavior. *Journal of Economic Psychology*, 41:1–11, 2014.

A. Bear and D. G. Rand. Intuition, deliberation, and the evolution of cooperation. *Proceedings of the National Academy of Sciences*, 113(4):936–941, 2016.

O. Bock, I. Baetge, and A. Nicklisch. hroot: Hamburg registration and organization online tool. *European Economic Review*, 71:117–120, 2014.

M. Buckert, J. Oechssler, and C. Schwieren. Imitation under stress. Technical report, WZB Discussion Paper, 2014.

A. Caplin and D. Martin. The dual-process drift diffusion model: Evidence from response times. *Economic Inquiry. doi: 10.1111/ecin.12294*, 2015.

A. W. Cappelen, U. H. Nielsen, B. Tungodden, J.-R. Tyran, and E. Wengström. Fairness is intuitive. *Experimental Economics*, pages 1–14, 2014.

V. Capraro and G. Cococcioni. Rethinking spontaneous giving: Extreme time pressure and ego-depletion favor self-regarding reactions. *Scientific reports*, 6:27219, 2016.

F. Chen and U. Fischbacher. Cognitive processes of distributional preferences: A response time study. *Research paper series Thurgauer Wirtschaftsinstitut*, 2015.

J. A. Clithero. Response times in economics: Looking through the lens of sequential sampling models. *Available at SSRN*, 2016.

A. Dreber, D. Fudenberg, D. K. Levine, and D. G. Rand. Altruism and self-control. *Available at SSRN 2477454*, 2014.

M. Drouvelis and B. Grosskopf. The effects of induced emotions on pro-social behaviour. *Journal of Public Economics*, 134:1–8, 2016.

S. Duffy and J. Smith. Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *Journal of Behavioral and Experimental Economics*, 51:47–56, 2014.

D. Engelmann and M. Strobel. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American economic review*, pages 857–869, 2004.

A. M. Evans, K. D. Dillon, and D. G. Rand. Fast but not intuitive, slow but not reflective: Decision conflict drives reaction times in social dilemmas. *Journal of Experimental Psychology: General*, 144(5):951, 2015.

E. Fehr and K. M. Schmidt. The economics of fairness, reciprocity and altruism experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1: 615–691, 2006.

S. Fiedler, A. Glöckner, A. Nicklisch, and S. Dickert. Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational behavior and human decision processes*, 120(2):272–284, 2013.

U. Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, 10(2):171–178, 2007.

S. Frederick. Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19 (4):25–42, 2005.

A. Hopfensitz and E. Reuben. The Importance of Emotions for the Effectiveness of Social Punishment. *The Economic Journal*, 119(540):1534–1559, 2009.

D. Kahneman. A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9):697, 2003.

M. G. Kocher, P. Martinsson, K. O. R. Myrseth, and C. E. Wollbrant. Strong, bold, and kind: Self-control and cooperation in social dilemmas. *Experimental Economics*, pages 1–26, 2012.

I. Krajbich, B. Oud, and E. Fehr. Benefits of neuroeconomic modeling: New policy interventions and predictors of preference. *The American Economic Review*, 104(5):501–506, 2014.

I. Krajbich, B. Bartling, T. Hare, and E. Fehr. Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature communications*, 6, 2015a.

I. Krajbich, T. Hare, B. Bartling, Y. Morishima, and E. Fehr. A common mechanism underlying food choice and social decisions. *PLoS Comput Biol*, 11(10):e1004371, 2015b.

J. Ledyard et al. Public goods: A survey of experimental research. Technical report, David K. Levine, 1997.

G. Loewenstein. Emotions in economic theory and economic behavior. *The American Economic Review*, 90(2):426–432, 2000.

G. Loewenstein and T. O Donoghue. The heat of the moment: Modeling interactions between affect and deliberation. *Unpublished manuscript*, 2007.

J. Lohse. Smart or selfish – When smart guys finish nice. *Journal of Behavioral and Experimental Economics*, 64(10):28–40, 2016.

J. Lohse, T. Goeschl, and J. H. Diederich. Giving is a Question of Time: Response Times and Contributions to an Environmental Public Good. *Environmental and Resource Economics*, pages 1–23, 2016.

G. Lotito, M. Migheli, and G. Ortona. Is cooperation instinctive? evidence from the response times in a public goods game. *Journal of Bioeconomics*, 15(2):123–133, 2013.

P. Martinsson, K. O. R. Myrseth, and C. Wollbrant. Social dilemmas: When self-control benefits cooperation. *Journal of Economic Psychology*, 45:213–236, 2014.

D. Mischkowski and A. Glöckner. Spontaneous cooperation for prosocials, but not for proselfs: Social value orientation moderates spontaneous cooperation behavior. *Scientific reports*, 6, 2016.

K. O. R. Myrseth and C. E. Wollbrant. Intuitive cooperation refuted: Commentary on Rand et al.(2012) and Rand et al.(2014). *University of Gothenburg WorkingPapers No 617*, 2015.

K. O. R. Myrseth and C. E. Wollbrant. Models inconsistent with altruism cannot explain the evolution of human cooperation. *Proceedings of the National Academy of Sciences*, 2016.

U. H. Nielsen, J.-R. Tyran, and E. Wengström. Second thoughts on free riding. *Economics Letters*, 122(2):136–139, 2014.

A. Nishi, N. A. Christakis, and D. G. Rand. Cooperation, decision time, and culture: Online experiments with american and indian participants. *Available at SSRN*, 2016.

J. Palmer, A. C. Huk, and M. N. Shadlen. The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of vision*, 5(5):1–1, 2005.

M. Piovesan and E. Wengström. Fast or fair? a study of response times. *Economics Letters*, 105 (2):193–196, 2009.

R. Polanía, I. Krajbich, M. Grueschow, and C. C. Ruff. Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron*, 82(3):709–720, 2014.

M. Rabin. Incorporating fairness into game theory and economics. *The American economic review*, pages 1281–1302, 1993.

D. G. Rand. Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science, Forthcoming*, 2016.

D. G. Rand, J. D. Greene, and M. A. Nowak. Spontaneous giving and calculated greed. *Nature*, 489(7416):427–430, 2012.

D. G. Rand, A. Peysakhovich, G. T. Kraft-Todd, G. E. Newman, O. Wurzbacher, M. A. Nowak, and J. D. Greene. Social heuristics shape intuitive cooperation. *Nature communications*, 5, 2014.

D. G. Rand, V. Brescoll, J. A. Everett, V. Capraro, and H. Barcelo. Social heuristics and social roles: Intuition favors altruism for women but not for men. *Forthcoming in Journal of Experimental Psychology: General*, 2016a.

D. G. Rand, V. L. Brescoll, J. A. Everett, V. Capraro, and H. Barcelo. Social heuristics and social roles: Intuition favors altruism for women but not for men. *Journal of Experimental Psychology: General*, 145(4):389, 2016b.

R. Ratcliff. Continuous versus discrete information processing: Modeling accumulation of partial information. *American Psychological Association*, 1988.

R. Ratcliff and J. N. Rouder. Modeling response times for two-choice decisions. *Psychological Science*, 9(5):347–356, 1998.

M. P. Recalde, A. Riedl, and L. Vesterlund. Error prone inference from response time: The case of intuitive generosity. 2014.

A. Rubinstein. Instinctive and cognitive reasoning: A study of response times*. *The Economic Journal*, 117(523):1243–1259, 2007.

L. Spiliopoulos and A. Ortmann. The bcd of response time analysis in experimental economics. *Available at SSRN 2401325*, 2015.

E. Stromland, S. Tjotta, and G. Torsvik. Cooperating, fast and slow: Testing the social heuristics hypothesis. 2016.

G. Tinghög, D. Andersson, C. Bonn, H. Böttiger, C. Josephson, G. Lundgren, D. Västfjäll, M. Kirchler, and M. Johannesson. Intuition and cooperation reconsidered. *Nature*, 498(7452): E1–E2, 2013.

G. Tinghög, D. Andersson, C. Bonn, M. Johannesson, M. Kirchler, L. Koppel, and D. Västfjäll. Intuition and moral decision-making–the effect of time pressure and cognitive load on moral judgment and altruistic behavior. *PloS one*, 11(10):e0164012, 2016.

P. Ubeda. The consistency of fairness rules: An experimental study. *Journal of Economic Psychology*, 41:88–100, 2014.

P. P. Verkoeijen and S. Bouwmeester. Does intuition cause cooperation? *PloS one*, 9(5):e96654, 2014.

P. Wright. The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of applied psychology*, 59(5):555, 1974.

Appendix A: Model and Predictions

In this part of the Appendix, we will describe the Drift Diffusion Model (DDM), introduced in section 2, in more detail. This description will illustrate how the DDM generates the following three predictions: First, the higher the subjective (absolute) utility difference between the options of choice, the lower a decision maker's expected response time. Second, a decision maker is more likely to make a mistake (i.e. to choose the option that yields a lower subjective utility) the smaller the utility difference between the two options of choice. Third, decision makers are more likely to make a mistake under time pressure if the subjective utility difference between the two options of choice is small.

The first two predictions are common in the DDM literature and have previously been used to show that the *correlation* between response times and "fair" behavior can reflect subjective utility differences (Krajbich et al., 2015a). The third prediction is novel - at least in the context of the literature on fairness and time pressure - and we will show that it follows immediately from the first and the second prediction under certain assumptions. The third prediction shows how time pressure affects behavior according to the DDM and, hence, provides insights into the interpretation of the overall effect of time pressure on choices which we discuss in Figure 1.

For the purpose of illustration, we will refer to a basic version of the DDM. Detailed descriptions of this model and potential extensions can be found e.g. in Ratcliff and Rouder (1998), Palmer et al. (2005) or Clithero (2016). This basic version of the DDM can be summarized as follows: A decision maker accumulates stochastic information about her preferences for a "fair" (henceforth: F) and a "selfish" (henceforth: S) option over a series of time periods t. We denote the decision maker's true underlying utility value for the fair and the selfish option by $u^i(F) = u_F$ and $u^i(S) = u_S$. Thus, the true underlying utility difference between the fair and the selfish option can be written as $V = u_F - u_S$. In each period t, decision makers observe noisy value signals $F_t \sim \mathcal{N}(u_F, \sigma_F^2)$ and $S_t \sim \mathcal{N}(u_S, \sigma_S^2)$ which are centered around the true means of the underlying value function. In each period, these signals are independently and identically distributed (i.i.d.). In line with the existing literature (Krajbich et al., 2015a), we will assume that $\sigma_F^2 = \sigma_S^2$, i.e. the distribution functions from which the signals are drawn only differ in their respective means. After observing a pair of signals, the decision maker computes the value difference between the two signals as $V_t = F_t - S_t$. The stochastic evidence observed until period t is then accumulated in a subjective state variable X_t. The accumulation process stops as soon as a state variable X_t crosses an upper threshold a^+, inducing the decision maker to choose F, or a lower threshold a^-, inducing the decision maker to choose S. We will assume that the two decision thresholds a^+ and a^- are equidistant from 0 so that $a^- = -a^+$. The evolution of the subjective state variable X_t can be written as:

$$X_t = X_{t-1} + d[V_t] = X_{t-1} + d[F_t - S_t] = X_{t-1} + d[u_F - u_S] + d[\epsilon_t] \tag{1}$$

where d represents the speed of the accumulation process and where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$ captures the noise in the process. For simplicity, we will set $d = 1$. Then, (1) can be re-written as

$$X_t = X_{t-1} + (u_F - u_S) + \epsilon_t = X_{t-1} + V + \epsilon_t \tag{2}$$

with

$$X_t \sim \mathcal{N}(tV, t\sigma^2) \tag{3}$$

For this simple variant of the DDM, expressions have been derived for the probability of choosing option F given that $V > 0$ and the mean response time given that $V \neq 0$ (Palmer et al., 2005; Clithero, 2016)[21]. Hence, this is the probability that a decision maker who prefers the fair option chooses the fair option. This probability can then be written as:

$$P_{F_F} = \frac{1}{1 + e^{\frac{-2Va_+}{\sigma^2}}} \tag{4}$$

The expected number of periods until one of the thresholds a^+ and a^- is reached for $V \neq 0$ can be written as[22]:

$$E[t] = \frac{a}{V} \tanh\left(\frac{aV}{\sigma^2}\right) \tag{5}$$

This is what is commonly referred to as "response time".[23]

From expressions (4) and (5) it is obvious that the probability of choosing the selfish option given that $V > 0$ (i.e. the probability that a fair decision maker chooses the selfish option by mistake) can be written as:

$$P_{S_F} = 1 - P_{F_F} \tag{6}$$

By symmetry equations (4)-(6) can be expressed equivalently for the probability that a selfish decision maker with $V < 0$ chooses the fair or the selfish option.

Prediction I

In this section we will demonstrate that a higher absolute utility difference between the fair and the selfish option will reduce the decision maker's expected response time. Since the first derivative of equation (5) w.r.t V is strictly negative for positive values of V the above statement follows immediately. Assuming symmetry, the same is true for negative values of V and the time it takes to reach the lower threshold.

Figure 4 illustrates how this prediction affects response times for two types of decision makers. Decision makers of type 1 prefer the fair option such that $V > 0$, whereas decision makers of type

[21] An additional assumption is that there is no initial bias in favor of one of the two options s.t. $X_0 = 0$. Given that the two thresholds are equidistant from 0, we replace a^+ and a^- by the absolute values.

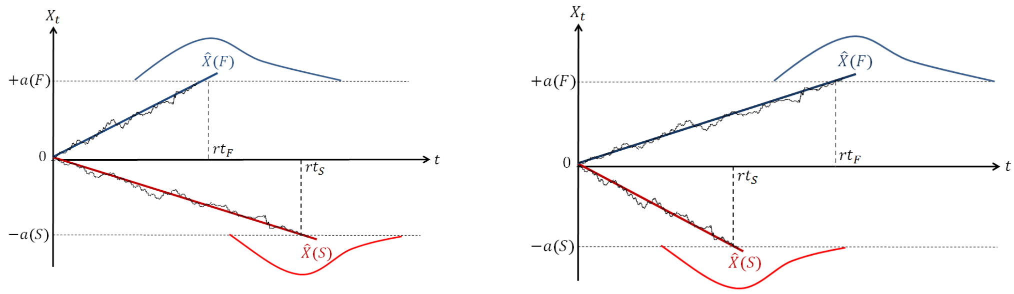
[22] This expression makes use of the hyperbolic tangent function $\tanh(z) = \frac{e^z + e^{-z}}{e^z - e^{-z}}$

[23] If the DDM is used to predict the actual response time in (mili)seconds, as it is typically done in the psychological literature, additional assumptions concerning the length of each period t are required. The statements made in this section are not confined to specific assumptions concerning the length of one period as we use the DDM as a model of decision making and not as an empirical model for estimating response times.

2 prefer the selfish option such that $V < 0$. We will denote the utility difference of the fair type 1 decision makers as V_F and that of the selfish type 2 decision makers as V_S. A direct implication of the first prediction is that if $V_F > |V_S|$, fair choices are relatively fast (see Figure 4a). Similarly, fair choices are expected to be relatively slow, if $V_F < |V_S|$ (see Figure 4b). This relationship between V and E[t] has been previously used to show that arbitrary correlations between fair choices and response times can be created by varying the relative attractiveness of the fair option (Krajbich et al., 2015a).

Figure 4: An illustration of two exemplary processes

(a) Fair choices are faster (b) Selfish choices are faster



Notes: Figure 1 displays two exemplary Drift Diffusion Process. $\hat{X}(F)$ ($\hat{X}(S)$) represents the expected evolution of the subjective state variable for the average decision maker with a preference for the fair (selfish) option. The expected response times are characterized as rt_F and rt_S. The actual evolution of the subjective state variable is subject to noise, as characterized by the black lines that fluctuate around $\hat{X}(F)$ and $\hat{X}(S)$. Above the upper decision threshold a^+ we display the expected response time distribution of fair choices. Below the lower threshold a^-, we display the expected response time distribution of selfish choices.

Prediction II

In this section we will show that a decision maker is less likely to make a mistake (i.e. to choose the option that yields the lower subjective utility value) as $|V|$ increases. We will first demonstrate that a decision maker is more likely to choose her truly preferred option as $|V|$ increases. For a decision maker who prefers the fair option ($V > 0$), this follows immediately from equations (4) and (6). The first derivative of statement (4) w.r.t. V is given by:

$$\frac{dP_{F_F}}{dV} = \frac{a^+ e^{-(a^+ V)}}{(1 + e^{-(a^+ V)})^2} > 0 \tag{7}$$

Hence, the probability of choosing the truly preferred fair option increases in V for $V > 0$. As a corollary, using statement (6), P_{S_F} decreases in V. By symmetry, an equivalent result can be derived for a decision maker who prefers the selfish option ($V < 0$). The intuition behind this prediction is that the noise in the decision process will have a larger weight on the value of the state variable when V is small.

Prediction III

In this section we will show how forcing the decision maker to take a decision within a pre-defined time limit affects the probability of choosing the preferred option. In particular, we demonstrate that under time pressure, a decision maker is more likely to choose her truly preferred option the larger $|V|$. Assume that the decision maker needs to decide within t_l periods. We assume that decision makers who have not reached the upper or lower decision threshold until t_l will choose randomly between the two options, i.e. decision makers are unaware of the signals which they have received so far and only receive information once either of the two thresholds is reached.[24] Thus, the frequency of mistakes under time pressure will depend on the probability of having reached one of the decision thresholds until t_l which, in turn, is a function of $|V|$.

First of all, note that the probability of reaching one of the thresholds until t_l is higher the larger $|V|$. This statement can be derived from statement (3), which shows that the probability of not having reached one of the thresholds can be written as:

$$P(X_{t_l}(V) \in [a^-, a^+]) = \int_{a^-}^{a^+} \frac{1}{\sqrt{2t_l\sigma^2}} e^{-\frac{1}{2}\left(\frac{x - V\cdot t_l}{2t_l\sigma^2}\right)^2} dx \tag{8}$$

Expression (8) illustrates that the probability of not having reached the threshold decreases in $|V|$. To show this formally, the first derivative of expression (8) has to be negative. We compute the derivative using the transformation rule and rewrite expression (8) as:

$$\frac{1}{\sqrt{2t_l\sigma^2}} \int_{\frac{a^- - V\cdot t_l}{2t_l\sigma^2}}^{\frac{a^+ - V\cdot t_l}{2t_l\sigma^2}} e^{-\frac{1}{2}(\widetilde{x})^2} d\widetilde{x} := P(V) \tag{9}$$

Then taking the first derivative w.r.t. to V yields:

$$\frac{1}{\sqrt{2t_l\sigma^2}} \left(-\frac{1}{2\sigma^2}\right) \left[e^{-\frac{1}{2}\left(\frac{a^+ - V\cdot t_l}{2t_l\sigma^2}\right)^2} - e^{-\frac{1}{2}\left(\frac{a^- - V\cdot t_l}{2t_l\sigma^2}\right)^2} \right] := P'(V) \tag{10}$$

$P'(V)$ is strictly negative if the expression within the square brackets is strictly positive, which holds if:

$$\left[e^{-\frac{1}{2}\left(\frac{a^+ - V\cdot t_l}{2t_l\sigma^2}\right)^2} > e^{-\frac{1}{2}\left(\frac{a^- - V\cdot t_l}{2t_l\sigma^2}\right)^2} \right] \tag{11}$$

Further simplifying this expression and exploiting the symmetry of barriers (i.e. $a^- = -a^+$) yields:

$$(a^+ - V \cdot t_l)^2 < (a^+ + V \cdot t_l)^2 \tag{12}$$

[24] A weaker assumption which leads to similar predictions would be that decision makers who are closer to either threshold have a higher probability of choosing the associated action.

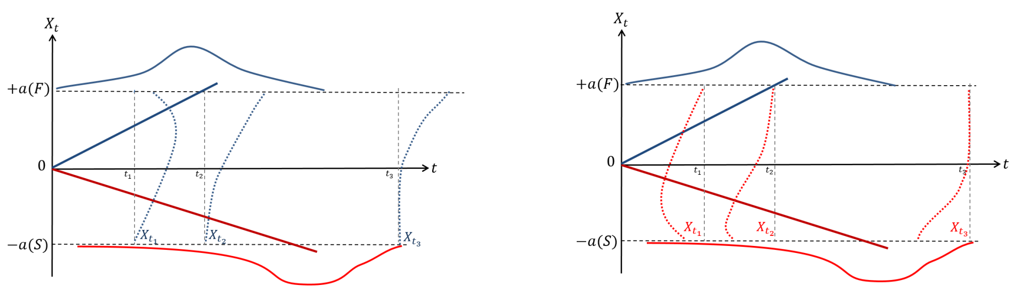
Obviously, this inequality is fulfilled iff $V \cdot t_l > 0$. By symmetry the same holds for a strictly negative V and a lower barrier $a^- = -a^+$.

A direct corollary, that follows from these expressions, is that stricter time limits will lead to more mistakes for a given $|V|$. A second corollary is that time pressure will increase the frequency of those choices which are associated with larger utility differences. To see this intuitively, assume that $V_F > |V_S|$, i.e. the utility difference is larger for decision makers who prefer the "fair" compared to decision makers who prefer the "selfish" option. Then a larger fraction of decision makers who prefer the "selfish" option will not have reached one of the decision thresholds at the point t_l and will, thus, choose randomly according to our assumption. As there are only two options available, half of these decision makers will choose the "fair" option by mistake.[25]

We illustrate the second corollary in Figures 2a and b. In both Figures, we depict a situation in which $V_F > |V_S|$ and hence "fair" choices are expected to be faster than "selfish" choices on average. We illustrate the effect of time pressure for three different time limits, t_1, t_2 und t_3. For each time limit, we depict the expected distribution of the state variable as $X_{t_i}, i = 1, 2, 3$. Trivially, time pressure only affects individuals who have not reached their decision threshold yet. For example, if the time limit was t_3, only few decision makers would actually face time pressure during their decision. Instead, if the time limit was t_1, most decision makers would face time pressure during their decision. We illustrate the role of utility differences for choice frequencies using time limit t_2. Figure 5a shows that the average decision maker who prefers the "fair" option has reached the upper threshold $+a$ until t_2. In contrast, only a comparatively small fraction of decision makers preferring the "selfish" option has reached the lower threshold $-a$ (see Figure 5b) until t_2. In other words, the fraction of decision makers who have not reached the threshold is larger for decision makers preferring the "selfish" option than for decision makers preferring the "fair" option. If all of these remaining decision makers choose randomly it is obvious that time pressure is more likely to induce mistakes among decision makers who have a weak preference for an option. In the case depicted here, forcing individuals to decide at t_2 would thus induce a higher frequency of fair choices compared to a baseline in which decision makers face no time limit (i.e. forcing decision makers to choose once t_3 has been reached).

Figure 5: An illustration of the link between time pressure and errors

(a) Decision makers who prefer the "fair" option (b) Decision makers who prefer the "selfish" option



More generally time pressure increases the fraction of those choices, which are associated with larger utility differences. Importantly, this means that in situations in which "fair" choices are subjectively less difficult than "selfish" choices, the "fairness is intuitive" hypothesis and the DDM

[25] Even if we relax the assumption of random choice and instead assume that decision makers will choose according to the sign of the state variable, it is still the case that the frequency of subjects whose state variable has the "wrong" sign (i.e. a positive sign if they prefer the selfish option) will be higher for smaller absolute values of V

both predict that time pressure should increase the fraction of "fair" choices relative to time delay. In contrast, the "fairness is intuitive" hypothesis and the DDM make opposite predictions whenever "fair" choices are subjectively more difficult than "selfish" choices.

Appendix B: Experimental Details

This part of the Appendix summarizes further details about the experiment. Part 1 shows how the games were displayed to the subjects in the experiment. All filler games used in the experiments are included in Part 2. The instructions are reproduced in Part 3.

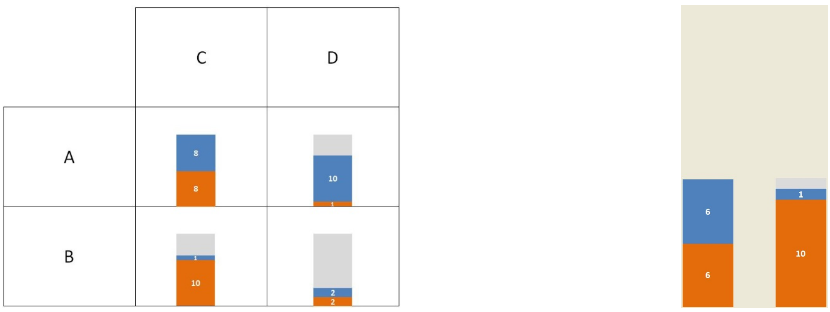
Visual Presentation of Games

The payoffs associated with each binary choice were displayed graphically as stacked and colored bars in all games (see Figure 6). The subject's own payoff corresponded to the orange and the other participant's payoff to the blue bar in all binary choice situations. Subjects received detailed instructions on how to read the bars and we confirmed their understanding in four control questions before they could start with the actual decision tasks. Furthermore, the examples used for the control questions did not relate to prisoner's dilemma games but displayed arbitrary payoffs to prevent potential priming effects. We believe that this way of displaying the payoffs has two advantages. First, it ensures that participants with different types of preferences can identify and implement their preferred choice with equal difficulty. Second, it makes the payoffs easily accessible and comparable across the choice options, even under response time constraints.

Figure 6: Presentation of games in experiment

(a) Prisoner's dilemma games

(b) Binary dictator games

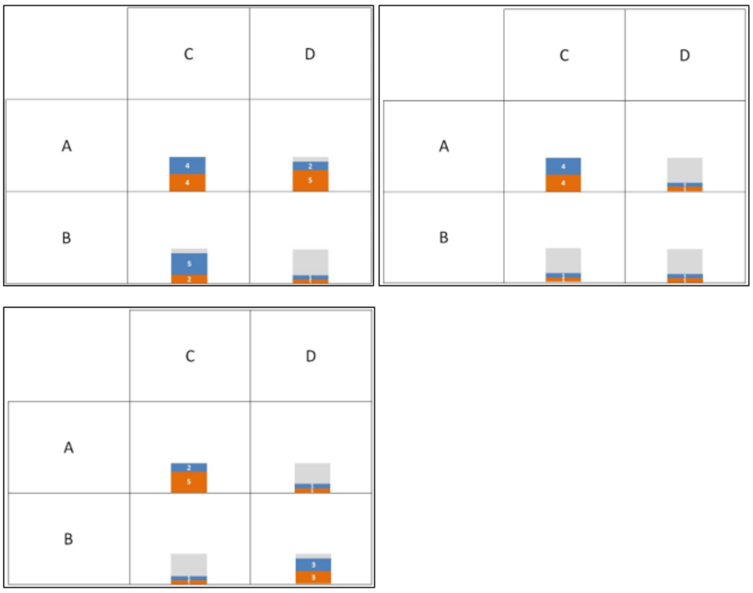


Filler Games

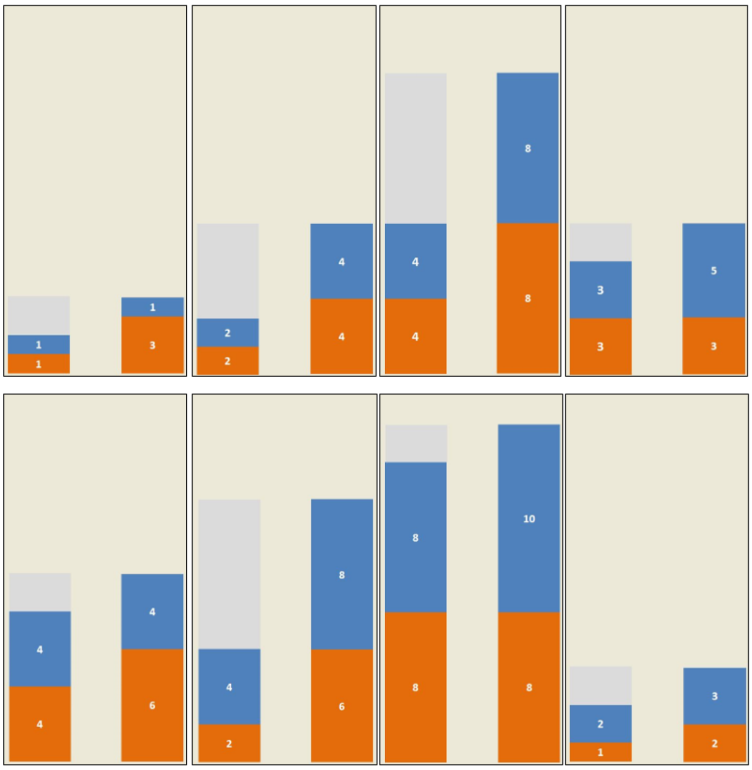
In Figure 7 we display the filler games that subjects faced during the experiment.

Figure 7: Filler Games

(a) Part 1 (Prisoner's Dilemma Games)



(b) Part 2 (Binary Dictator Games)



Instructions

Instructions were presented on the screens in German language. A translated version of the original instructions is presented below. The original instructions are available upon request.

Instructions for part 1 of the experiment

You will now start with the first part of the experiment.

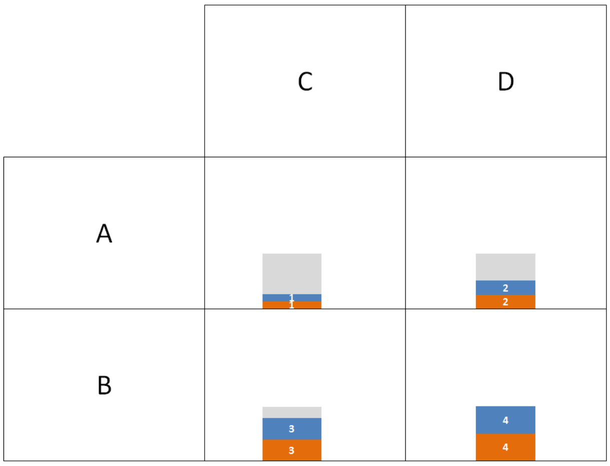
This part of the experiment consists of **10 rounds**.

In each round, you will interact with one other randomly chosen participant. No participant is going to be informed with whom he or she has interacted during the experiment.

Procedure within each round

In each round, both participants **simultaneously** choose one of two options: You decide between option “A“ and “B“, the other participant decides between option “C“ and “D“. Hence, you decide between options A and B without knowing which option has been chosen by the other participant.

Your payoff and the payoff of the other participant depend on the decisions of both participants. At the beginning of each round (so before you and the other participant have made a choice), both participants will see a table in which the four different payoffs are displayed.



Each of the four possible payoffs is depicted as a bar chart. The bars consist of several coloured parts.

Your own payoff corresponds to the orange part of the bar. The number within the orange part of the bar indicates the exact Euro amount that you will going to receive in that case.
The payoff of the other participant corresponds to the blue part of the bar. The number within the blue part of the bar indicates the exact Euro amount that the other participant is going to receive in that case.

The height of the orange and the blue part corresponds to the sum of payoffs for both participants. The grey part of the bar indicates the payoff difference to the maximum achievable sum in this round.

Examples:

Example 1:

You have chosen option A, the other participant has chosen option C. This results in the following payoffs: You receive 1 Euro and the other participant receives 1 Euro.

Example 2:

You have chosen option B, the other participant has chosen option D. This results in the following payoffs: You receive 4 Euro and the other participant receives 4 Euro.

Please note

The actual payoff table is going to **look different** in the experiment. Also, the payoffs will differ in each round.

End of a round

The round is over as soon as both participant have taken a decision. You **will not be informed** about the choices of the other participant.

Your payoff

At the end of the experiment the computer will randomly **choose one round from this or the other part of the experiment. You receive the amount which results from your own and the decision of the other participant.** Hence, each decision in this part of the experiment can influence your final payoff at the end of the experiment

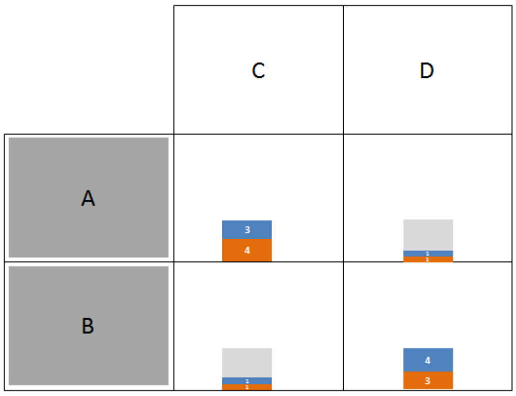
You have received all instructions for the first part of the experiment now. Press "continue" to learn more about how each choice will be displayed on your computer and to test your comprehension on an example.

Screen:

Example

In this part you can test your comprehension using the payoff table displayed below. Your choices in this part will not influence your final payoff.

Please look at the payoffs displayed in the table:



Question 1:

Suppose you choose option A and the other participant chooses option C.

Your payoff in Euro: ____________________
Payoff of the other participant in Euro: ___________

Question 2:

Suppose, you choose option B and the other participant chooses option C.

Your payoff in Euro: ____________________ Euro (1)
Payoff of the other participant in Euro: ___________ Euro (1)

Press “continue“ to find out whether you answered correctly.

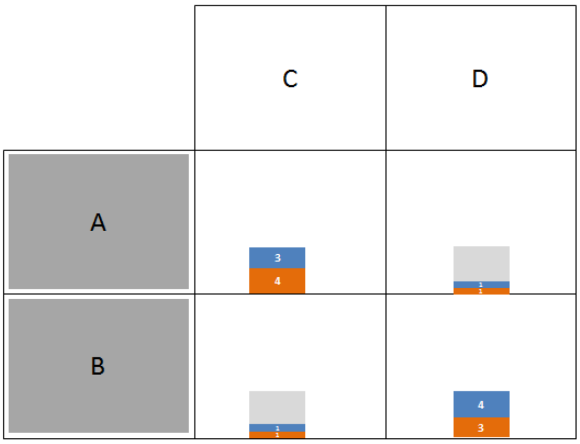
<u>Screen:</u>

<u>Feedback (correct):</u>

You have answered both questions correctly. You can start with round 1 now.

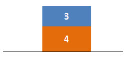
<u>Feedback (wrong):</u>

Unfortunately, you did not answer all questions correctly. Please take a look at the payoffs displayed in the table once more:



<u>Question 1 (wrong):</u>

In question 1 you were asked: What payoff would you and the other participant receive in the case that you would chose option A and the other participant would choose option C. The payoffs are as follows:



In this case you would receive 4 Euro (orange part of the bar). The other participant would receive 3 Euro (blue part of the bar).

Please make sure that you have understood all instructions. If you need further help please contact the experimenter.

<u>Question 2 (wrong):</u>

In question 2 you were asked: What payoff would you and the other participant receive in the case that you would chose option B and the other participant would choose option C. The payoffs are as follows:

In this case you would receive a payoff of 1 Euro (orange part of the bar). The other participant would receive a payoff of 1 Euro (blue part of the bar).

Please make sure that you have understood all instructions correctly. If you still have problems, please contact the experimenter.

Screen:

[Time Pressure]

In the following 5 rounds you should decide quickly.

Please select option A or B in **less than 12 seconds** in every round.

The remaining decision time is displayed above the payoff table.

If your decision takes longer than 12 seconds in one round and if this round is chosen for payoff, you will not receive your show-up fee of 3 Euro.

[Time Delay]

In the following 5 rounds you should wait before making a decision.

In each round you should wait **at least 12** seconds before you decide between options A und B.

Only after 12 seconds have passed, the grey choice buttons labelled “A“ and “B“ will appear on the screen.

You don’t have to decide precisely after 12 seconds. You **can think as long as you want**.

Press “continue“ to start with the first round.

Screen:

[Time Pressure]

In the following 5 rounds, there is a **time limit** for your decision.

Suppose, one round from this part of the experiment is chosen. **You receive your show up fee of 3 Euro only if you …**

… took a decision in more than 12 seconds
… took a decision in less than 12 seconds
… took a decision in exactly 12 seconds
… took a decision in less than 20, but more than 12 seconds

… in the randomly chosen round.

Please select the correct answer. On the next screen you will be informed whether your answer was correct.

[Time Delay]

In the following 5 rounds there is a **time limit for your decision.**

In each round you should…

… take a decision in less than 12 seconds.
… wait at least 12 seconds before to take a decision.
… take a decision in exactly 12 seconds
… take a decision in at least 8, but less than <TimePressure|1> seconds to take a decision.

Screen:

Feedback Correct:

You answered the question correctly and can now start with round 1.

Feedback Wrong:

Unfortunately, you haven't answered the question correctly. Please take a look at the following advice again.

[Time Pressure]

In the following 5 rounds you should decide quickly.

Please select option A or B in **less than 12 seconds** in every round.

The remaining decision time is displayed above the payoff table.

If your decision takes longer than 12 seconds in one round and if this round is chosen for payoff, you will not receive your show-up fee of 3 Euro.

[Time Delay]

In the following 5 rounds you should wait before making a decision.

In each round you should wait **at least 12** seconds before you decide between options A and B.

Only after 12 seconds have passed, the grey choice buttons labelled "A" and "B" will appear on the screen.

You don't have to decide precisely after 12 seconds. You **can think as long as you want**.

Press „continue" in order to start with the first round.