

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

Anna Louisa Merkel¹  · Johannes Lohse²

Received: 12 December 2016 / Revised: 2 November 2017 / Accepted: 12 March 2018 /
Published online: 22 March 2018
© Economic Science Association 2018

Abstract Evidence from response time studies and time pressure experiments has led several authors to conclude that “fairness is intuitive”. In light of conflicting findings, we provide theoretical arguments showing under which conditions an increase in “fairness” due to time pressure indeed provides unambiguous evidence in favor of the “fairness is intuitive” hypothesis. Drawing on recent applications of the Drift Diffusion Model (Krajbich et al. in *Nat Commun* 6:7455, 2015a), we demonstrate how the subjective difficulty of making a choice affects decisions 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 binary dictator and prisoner’s dilemma games under time pressure or time delay. In addition, we manipulate the subjective difficulty of choosing the fair relative to the selfish option. Our main finding is that time pressure does not consistently promote fairness in situations where this would be predicted after accounting for choice difficulty. Hence, our results cast doubt on the hypothesis that “fairness is intuitive”.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10683-018-9566-3>) contains supplementary material, which is available to authorized users.

✉ Anna Louisa Merkel
anna.merkel@awi.uni-heidelberg.de

Johannes Lohse
j.lohse@bham.ac.uk

¹ Department of Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany

² Department of Economics, University of Birmingham, J G Smith Building, Edgbaston Campus, Birmingham, UK

Keywords Distributional preferences · Cooperation · Time pressure · Response times · Cognitive processes · Drift Diffusion Models

JEL Classification C32 · C72 · C91 · D64 · D91 · H41

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 “fairness” is intuitive and automatic or follows from a deliberative weighting of the costs and benefits of making a fair choice. Several authors have approached this question by analyzing response times as a proxy for deliberation (Rubinstein 2007; Spiliopoulos and Ortmann 2017). A popular method for understanding the causal impact of deliberation on choices is to place subjects under time pressure or time delay, given that subjects who are constrained to make a fast choice might increase their reliance on intuition compared to subjects who are constrained to wait before making a choice (Wright 1974; Rand et al. 2012).

Using this method, 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. These results have inspired the “fairness is intuitive” (FII) (Cappelen et al. 2016) hypothesis. According to the FII hypothesis, a decision maker intuitively prefers fairness, i.e. cooperation in a public good or sharing resources in a dictator game.¹ However, this predisposition towards fairness can be overridden by a more deliberative weighting of the costs and benefits, such that deliberation can promote selfishness (Rand et al. 2012).

The FII hypothesis has not been unequivocally confirmed empirically. In contrast to the original results of Rand et al. (2012), Tinghög et al. (2013), Verkoeijen and Bouwmeester (2014), and Bouwmeester et al. (2017) do not find that constraining deliberation by time pressure increases the fraction of cooperative choices in one-shot public good games. Furthermore, Tinghög et al. (2016) find that time pressure does not affect the fraction of fair choices in (modified) dictator games. Finally, findings in Capraro and Cococcioni (2016) and Lohse (2016) suggest that placing subjects under stronger time pressure leads to more selfish choices in public good games. Similarly, Mrkva (2017) finds that time pressure leads to an increase of selfish choices in modified dictator games with high stakes, but not with low stakes.

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

¹ Obviously, the economics literature has come up with various notions of what constitutes a “fair” choice (Rabin 1993; Engelmann and Strobel 2004; Fehr and Schmidt 2006). In Sect. 2, we will describe in more detail what we refer to as a “fair” choice in the context of our paper.

deliberation also affect response times and thereby distort the identification of intuitive or deliberative choices from fast and slow responses (Recalde et al. 2014; Krajbich et al. 2015a). We explore how the subjective difficulty of choosing between a fair and a selfish option, as one such factor, affects choices under time pressure and time delay. Our theoretical predictions highlight that, without controlling for the effect of choice difficulty in a given situation, observing a positive effect of time pressure is not necessarily evidence in favor of the FII hypothesis; and observing no effect of time pressure is not necessarily evidence against the FII hypothesis. Thereby, we provide one plausible explanation why previous tests of the FII hypothesis might have come to different conclusions.

Our theoretical considerations are based on insights from a recent paper by Krajbich et al. (2015a) who use a Drift Diffusion Model (DDM) to illustrate how choice difficulty may affect response times. The central prediction of the DDM is that more difficult choices, i.e. those in which the utility difference between the fair and the selfish option are small, are associated with longer response times. We build on this insight and explore how the subjective difficulty of making a choice affects a *causal* test of the FII hypothesis. Our analysis is based on the assumption that choices under time pressure may be affected by both, the amount of deliberation involved in the choice and the subjective difficulty of making a choice (Alós-Ferrer 2016). Hence, the overall effect depends on how time pressure affects choices according to the FII hypothesis *as well as* the DDM.

We use a simple version of the DDM to show that time pressure causes decision makers who perceive smaller utility differences to make more mistakes. Thus, the DDM predicts that time pressure can either increase the fraction of fair choices, if fair decision makers perceive larger utility differences and are less common in the population; or decrease the fraction of fair choices, if selfish decision makers perceive larger utility differences and are less common in the population. The mechanism motivating the FII hypothesis, on the other hand, predicts that time pressure should always increase the fraction of fair choices in one-shot games. Hence, whenever fair decision makers perceive larger utility differences than selfish decision makers and are less common in the population, the DDM and the FII both predict that time pressure should increase the fraction of fair choices. Observing a positive effect of time pressure in these situations can therefore only provide ambiguous evidence in favor of the FII hypothesis as the same pattern could also be explained by the DDM, while observing no effect is unambiguous evidence against the claim that “fairness is intuitive”. On the other hand, whenever selfish decision makers perceive larger utility differences than fair decision makers and are less common in the population, the FII hypothesis and the DDM predict opposite effects of time pressure, which may even cancel each other out. Observing no or even a negative effect of time pressure in these situations is not sufficient to unambiguously reject the claim that “fairness is intuitive”, while observing a positive effect provides unambiguous evidence in favor of the FII hypothesis. These arguments illustrate that classifying a choice situation into one of these types is central for the correct interpretation of time pressure effects. Otherwise the FII hypothesis could be spuriously accepted or rejected. The fact that previous studies have not explicitly

accounted for subjective utility differences might explain why they have arrived at different conclusions.

To causally test the FII hypothesis, we conduct an experiment in which subjects take decisions under time pressure or time delay in multiple two-person binary dictator and prisoner's dilemma games. Across games, 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 such that the DDM and the FII make either consistent or opposite predictions concerning the effect of time pressure. To classify choice situations into one of these two possible types, we use an additional treatment, in which subjects are unconstrained in their response time and in which we observe response time correlations and choice frequencies. 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.

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 in each game, once under time pressure and thereafter under time delay. Second, by comparing evidence from binary dictator and prisoner's dilemma games, we are able to distinguish between fair choices in non-strategic and strategic decisions. Thereby, we investigate whether pro-social behavior follows a common cognitive pattern across different contexts. While several previous tests of the FII hypothesis are based on evidence from strategic decisions in public good or prisoner's dilemma games, non-strategic decisions in simple binary dictator games might allow for a more direct test given that they are unconfounded by strategic uncertainty or misconceptions regarding the game.

Overall, our analysis provides at most limited empirical support for the hypothesis that "fairness is intuitive". In those binary dictator and prisoner's dilemma games, in which our classification suggests that time pressure should increase fairness according to both models, we do not observe such an increase across all between-subjects tests. In the same games, there is no consistent within-subjects evidence that subjects who choose the fair option under time pressure are more likely to switch to the selfish option under time delay. In binary dictator games in which an increase of fair behavior under time pressure would constitute unambiguous evidence in favor of the FII hypothesis, we do not find that time pressure significantly increases the frequency of fair choices. This evidence holds between- and within-subjects. A complementary analysis shows that switching patterns strongly reflect choice difficulty (subjective indifference), a pattern that is supported by the DDM but not by the FFI hypothesis.

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

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 follows a cooperation heuristic that individuals have developed in repeated everyday interactions. Upon deliberation, the same individuals may realize that there are no strategic incentives to cooperate in atypical one-shot situations implemented in the lab which leads to more defection. Cooperation is the most prominent application of the “Social Heuristics Hypothesis”. Its underlying mechanism could, however, apply more broadly to non-strategic choices in the dictator game assuming that sharing resources with other people is also an advantageous long-term strategy because of reciprocity or reputation concerns. We summarize the claim that intuition promotes fairness across different contexts as the FII hypothesis.

The FII hypothesis generates empirically testable predictions concerning the effect of time pressure and time delay on fairness. Since heuristics are seen to operate relatively independently from the details of a choice situation, the FII hypothesis predicts that the same decision maker should be more likely to choose the fair option when placed under time pressure than when she makes a deliberative choice. Similarly, when observing choices of different decision makers, subjects who are placed under time pressure should on average choose the fair option more frequently than subjects constrained to wait before making a choice.

However, the observation that time pressure leads the same decision maker to choose the fair option with higher probability or that time pressure increases the fraction of fair choices cannot be unambiguously interpreted as evidence in favor of the FII hypothesis without accounting for choice difficulty. 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 describe the DDM in more detail.²

2.1 Time pressure in the Drift Diffusion Model (DDM)

Assume that a single decision maker faces a binary choice between a “fair” (F) and a “selfish” (S) option. According to the DDM, this decision maker is initially unaware of the utility value she receives from these options. However, she can accumulate stochastic information regarding her preferences 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 normally distributed around her true underlying utility values. The difference between the two signals $F_t - S_t$ is added to a subjective state variable X_t^i which, thus, encodes the probability that F yields a higher utility than S (Krajbich

² 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).

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 b , inducing the decision maker to choose S . The length of the accumulation process, i.e. the number of time periods before the upper or lower threshold is reached, corresponds to the decision maker's response time.

The standard DDM makes two predictions regarding the theoretical distribution of response times and decision errors (e.g., Ratcliff and Rouder 1998).³ First, the decision maker's response time depends on the underlying absolute 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 absolute utility difference is small because she has to sample fewer signals to reach one of the thresholds. Second, given that the final decision is reached by observing a series of noisy signals, the decision maker is 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 two options. 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 decision makers with otherwise longer response times to make a choice before being sufficiently sure about their truly preferred option. Thus, time pressure is equivalent to a reduction in the decision thresholds. This induces decision makers to choose at lower precision because noise has a higher likelihood of influencing their decision. Importantly, the likelihood of making a mistake is larger for decision makers with smaller absolute utility differences because their value signals contain relatively less information relative to noise.

Aggregating these individual level effects provides predictions for how overall choice frequencies are affected by time pressure. For illustrative purposes, we will distinguish between three situations, labeled *type 0*, *type 1* or *type 2*. Furthermore, we will refer to a decision maker as "selfish" or "fair" depending on which of the two options yields a higher utility value according to her subjective preferences. In situations of *type 0*, the incentives are such that the underlying absolute utility differences are the same for the average selfish and fair decision maker. Thus, fair and selfish decision makers are equally likely to make a mistake under time pressure and time delay. In situations of *type 1*, on the other hand, the absolute utility difference is larger for the average fair than for the average selfish decision maker. Hence, in these situations selfish decision makers are more likely to make a mistake. Finally, in situations of *type 2*, the utility differences are larger for the average selfish than for the average fair decision maker such that fair decision makers are more likely to make a mistake.

³ A more detailed description of the DDM as well as proofs and derivations of all predictions are contained in Online Appendix A.

Under the simplifying assumption that time pressure exclusively affects decision makers with smaller average utility differences (i.e. weak preferences for one of the options) and that there are no mistakes under time delay, the DDM generates straightforward predictions. In situations of *type 1*, time pressure exclusively causes selfish decision makers to make a mistake such that time pressure inflates the frequency of fair choices relative to a situation without time pressure. For situations of *type 2*, the DDM predicts the reverse effect. Here, fair decision makers should make more mistakes under time pressure, thus reducing the fraction of fair choices under time pressure.

Without this simplifying assumption (i.e. assuming that the probability of making a mistake is positive under time pressure and, to a smaller degree, under time delay for all decision makers), the DDM predictions depend on two factors: first, the average strength of preferences and second, the relative frequency of fair and selfish decision makers within the population.⁴ The strength of preferences determines the likelihood of committing an error under time pressure and time delay for a given type of decision maker. The population shares, on the other hand, determine the resulting absolute number of mistakes and the aggregate direction of switches. The most simple test case for the FII hypothesis is a situation of *type 0* in which the relative population shares of fair and selfish decision makers are roughly similar. In such a perfectly balanced situation—however infrequently such situations might occur in actual empirical tests—the DDM predicts that time pressure should have no effect on the frequency of fair choices since fair and selfish decision makers are equally likely to make a mistake (under time pressure and time delay) and both groups are of equal size. Consequently, the absolute number of mistakes is perfectly balanced between both groups and there should be no effect of time pressure. The DDM also generates unambiguous predictions when the type of decision maker, who has larger utility differences, is less common within the population (< 50%). In these cases, the DDM predicts that time pressure increases the fraction of choices which are associated with larger absolute utility differences. For example, if the fair option is preferred by less than 50% of subjects in a situation of *type 1* (where fairness is “easy”), time pressure should increase the fraction of fair choices. This increase is driven by two factors: first, selfish decision makers are more likely to make an error under time pressure and to switch to their preferred choice under time delay as compared to fair decision makers. Second, given that they constitute the larger group, there should be more switches from the fair (under time pressure) to the selfish option (under time delay) than vice versa.

In all other cases, i.e. when the decision makers who have larger utility differences are more common in the population, the predictions of the DDM depend on the relative population shares as well as the unobservable difference in error rates under time pressure and time delay for both types of decision makers.⁵

⁴ We are grateful to an anonymous referee for pointing out this crucial distinction and helping us to refine our model.

⁵ Online Appendix A contains a more formal discussion of the possible results.

Table 1 Testing the FII hypothesis

	Predicted effects of time pressure		
	Type 0 $p(f) = 0.5$	Type 1 $p(f) < 0.5$	Type 2 $p(f) > 0.5$
Observed effect			
↑	Unambiguous accept (0a)	Ambiguous accept (1a)	Unambiguous accept (2a)
↔	Unambiguous reject (0b)	Unambiguous reject (1b)	Ambiguous accept (2b)
↓			

2.2 Testing the FII hypothesis accounting for DDM predictions

Assuming that choices under time pressure and time delay are affected by the relative use of intuition over deliberation *as well as* the subjective difficulty of making a choice, the arguments above imply that the predictions of the DDM and the FII are congruent in situations of *type 1* as long as the fraction of fair decision makers is smaller than 50%. Hence, observing that time pressure increases the fraction of fair choices in these situations 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). Instead, if we do not find these predicted patterns, then this constitutes unambiguous evidence against the FII hypothesis (1b).⁶

In contrast, unambiguous evidence in favor of the FII hypothesis can be obtained from situations of *type 2*, as long as the fraction of selfish decision makers is smaller than 50%. Here, the FII hypothesis and the DDM predict opposite time pressure effects which may even cancel each other out. Thus, observing that time pressure does increase the fraction of fair choices would be unambiguous evidence in favor of the FII hypothesis (2a). Not observing any or even a negative effect would not necessarily be inconsistent with the FII hypothesis because the opposite effects of the FII hypothesis and the DDM may actually cancel each other out (2b).

Finally, in situations of *type 0* in which relative population shares are roughly similar, the DDM should have little influence on the direction of time pressure effects as fair and selfish decision makers are equally likely to make mistakes and are present in equal proportions within the population. Thus, observing an increase of fair behavior in such situations would be unambiguous evidence in favor of the FII, while observing no or a negative effect would provide unambiguous evidence against the FII.

Whenever the DDM predictions regarding the direction of time pressure effects are not clear because they depend on unobservable differences in error rates, tests of the FII hypothesis cannot be interpreted unambiguously. Therefore, classifying the

⁶ Note that observing no effect is not necessarily evidence against the DDM in these situations. This is because the true model might be that “selfishness is intuitive”. Hence, a negative effect of time pressure attributable to the “selfishness is intuitive” model might be cancelled out by a positive effect attributable to the DDM. For this reason, we cannot *jointly* reject the FII hypothesis *and* the DDM.

choice situation as *type 0*, *type 1* or *type 2* and approximating the population shares of fair decision makers is necessary for correctly interpreting the evidence. Previous tests of the FII hypothesis might, thus, suffer from spuriously accepting the FII hypothesis based on observing an increase of fairness in situations of *type 1* or spuriously rejecting it based on observing no effect or a decrease of fairness in situations of *type 2*.

3 Experimental design

In our experiment, we collect decisions from four binary dictator (see Table 2) and four prisoner's dilemma games (see Table 3). In each game, subjects are asked to choose between a "fair" and a "selfish" option (labeled option "A" or "B" on the decision screen). In line with the FII hypothesis (Rand et al. 2014), we label a choice as "fair" if it implies sharing resources with another individual at own costs. According to this definition the equal allocation is the "fair" choice in the binary dictator (BD) games and cooperation is the "fair" choice in the prisoner's dilemma (PD) games. Across the four BD and PD games, we increased the social benefits of choosing the fair option from VERY LOW to HIGH. 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.⁷

If subjective utility differences reflect the costs and benefits of choosing the fair option (Andreoni and Miller 2002), we would expect that fair decision makers perceive smaller utility differences in the VERY LOW games than selfish decision makers. 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 games potentially resemble a type 2 choice situation that would allow for an unambiguous test of the FII hypothesis. By the same logic, we expect that the HIGH games resemble a type 1 choice situation in which fair decision makers perceive larger utility differences than selfish decision makers. 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.

Despite these considerations, it is hard to predict a priori if choosing the fair option will be subjectively more or less difficult than choosing the selfish option in a given game. Furthermore, a correct interpretation of the evidence also requires a measure of whether the fair or the selfish option is preferred by a majority of decision makers. To gain empirical insights into the subjective difficulty of choosing the fair and the selfish option as well as the respective population shares, we conducted additional sessions in which subjects could decide without being constrained in their response times. Based on the previous finding that response

⁷ Labeling the equal outcome as fair in the binary dictator game also aligns our FII predictions with recent findings in Capraro et al. (2017) who show that equal outcomes are preferred by intuitive decision makers whereas deliberation allows for a variety of motives to affect decisions.

Table 2 Binary dictator games used in the experiment

VERY LOW		LOW		MEDIUM		HIGH	
Unequal	11, 0	Unequal	9, 0	Unequal	10, 1	Unequal	15, 2
Equal	1, 1	Equal	3, 3	Equal	6, 6	Equal	11, 11

Table 3 Prisoner's dilemma games used in the experiment

VERY LOW		LOW		MEDIUM		HIGH					
C	D	C	D	C	D	C	D				
C	3.10, 3.10	1, 5.10	C	4, 4	1, 6	C	6, 6	1, 8	C	8, 8	1, 10
D	5.10, 1	2, 2	D	6, 1	2, 2	D	8, 1	2, 2	D	10, 1	2, 2

times reflect the relative difficulty of the choice situation (Krajbich et al. 2015a), we use these additional observations to classify games as type 0, 1 or 2.

We used the following procedures in our experiment: Part 1 of the experiment consisted of two successive blocks. In block 1, subjects made decisions in the four prisoner's dilemma games displayed in Table 3 in randomized order. After each prisoner's dilemma game, subjects made choices in unrelated filler games (see Online Appendix B). Once subjects had completed block 1 and a short questionnaire, we elicit choices in 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.⁸

Part 2 of the experiment also consisted of two successive blocks. In block 1, subjects made choices in the four binary dictator games displayed in Table 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 Online Appendix B). Once subjects had completed block 1 and another short questionnaire, they made choices in the same four binary dictator and filler games again in block 2.

For each binary choice, subjects were randomly re-matched in pairs and no feedback on their partner's choice was given until the very end of the experiment. At the end of the experiment, one of the games was randomly drawn and subjects were paid according to their own and their partner's choice.

To analyze the effect of time pressure on the fraction of fair choices, we randomly assigned subjects to one of four (between-subjects) conditions, in which we implemented different response time constraints: in the two Time Pressure

⁸ 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. Subjects were not informed that they would make the same choices in both blocks.

conditions, TP and STP, subjects were constrained to choose under time pressure in block 1 and forced to wait before making a choice in block 2. In the Time Delay (TD) condition, subjects were forced to delay their decision in both, block 1 and block 2. In the Unconstrained condition subjects did not face an exogenous time limit in either block.

Our experimental design is summarized in Table 4. In the Time Pressure (TP) condition, the time limit was 12 s for all prisoner's dilemma games and 6 s for all binary dictator games. These time limits correspond to the first quartile of the response time distribution of the *first* choice in the Unconstrained condition.⁹ Given that subjects usually get faster over time and that it is unclear how much time is required to induce intuitive decisions¹⁰, we implemented a stricter time limit of 8 s in the PD games which was reduced to 4 s in the BD games in the Strong Time Pressure (STP) condition. These time limits correspond to the first quartile of the response time distribution for the *last* decision in the Unconstrained condition. The time delay limit was 12 s for the PD games and 8 s for the BD games in both the TP and the STP condition, so that there is a small gap in the STP condition. The payoffs were displayed graphically as stacked and colored bars in all games (see Online Appendix B) in order to make them easily accessible and comparable, even under time pressure.

To ensure compliance with our treatment, we forced subjects to delay their decision by displaying the choice buttons only after 12 s (6 s) 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.¹¹ A counter, displaying seconds spent, was included on each decision screen in both the Time Delay and the Time Pressure conditions.

At the end of part 1, we elicited subjects' beliefs regarding the choices of other participants which allows us to test whether time pressure and time delay affected beliefs differently.¹² Subjects were paid an additional Euro for a correct guess. In addition, we asked subjects to provide a subjective assessment regarding which of the two options they perceive as the fairer option for the very first BD and PD games

⁹ 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 s which corresponds to the median response time in their response time correlation study. Buckert et al. (2017) define time pressure as 2/3 of the median response time in a Cournot game. Our analysis of response times in the Unconstrained treatment 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.

¹⁰ For instance, Myrseth and Wollbrant (2017) argue that any time limit above 4 s could allow decision makers to engage in some level of deliberation. Spiliopoulos and Ortmann (2017) report that mean response times fall by up to 30% when subjects face the exact same game multiple times.

¹¹ 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; Bouwmeester et al. 2017). In contrast, we observed few violations of the time limit: Averaged over all decisions and treatments, the time pressure conditions were violated in 2.5% of the BD and 1.7% of the PD games. There is no significant difference in violations between the TP and STP condition.

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

Table 4 Experimental design

	Time pressure (TP)	Strong Time Pressure (STP)	Time delay (TD)	Unconstrained (U)
Part 1				
Block 1	≤ 12 s	≤ 8 s	> 12 s	No constraint
4 PDs + 4 filler games				
Block 2	> 12 s	> 12 s	> 12 s	No constraint
4 PDs + 4 filler games				
Part 2				
Block 1	≤ 6 s	≤ 4 s	> 6 s	No constraint
4 BDs +12 filler games				
Block 2	> 6 s	> 6 s	> 6 s	No constraint
4 BDs + 12 filler games				

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"

they encountered in each block. This assessment can be used to identify if the equal (cooperative) option is indeed perceived as "fair" by a majority of our subjects.¹³

4 Results

The experiment was conducted at the University of Heidelberg AWI Lab. In total, 238 undergraduate and graduate students of all disciplines were recruited to participate in the experiment (62 in Unconstrained, 74 in Time Delay, 72 in Time Pressure and 30 in Strong Time Pressure) 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 Online Appendix B) on the screen and questions were answered privately. At the end of the experiment, subjects were paid in private. The average earnings were 12 Euro, including a 3 Euro show-up fee. In the following, we will report the results of the Unconstrained condition before analyzing the results of the Time Pressure and Time Delay conditions.

¹³ Despite being unincentivized and thus noisy, this survey approach can provide some insights into the modal fairness perceptions of subjects (Faravelli 2007; Cubitt et al. 2011; Reuben and Riedl 2013).

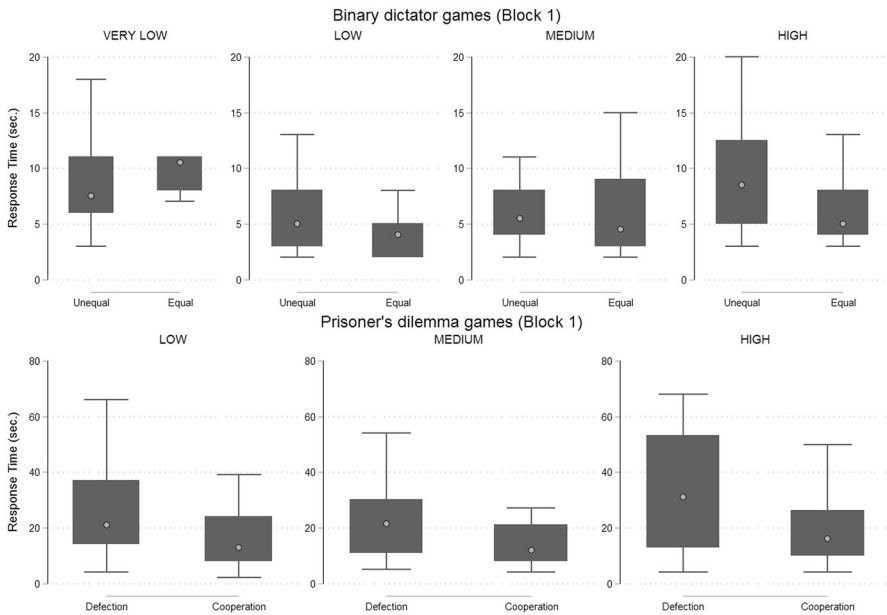


Fig. 1 Response times in prisoner's dilemma and binary dictator games (unconstrained condition)

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 approximate the frequency of fair and selfish choices. This information can be used to classify the different games according to the theoretical considerations outlined in Sect. 2.

In Fig. 1 (top panel) we compare the distribution of response times between choices of the equal (“fair”) and the unequal (“selfish”) option in the BD games. The frequency of “fair” choices rises significantly from 9.7% in the VERY LOW game to 43.5% in the LOW, 51.6% in the MEDIUM, and to 61.3% in the HIGH game (Pairwise Sign Test, $p < 0.01$).¹⁴ In line with the results in Krajbich et al. (2015a), we observe that the correlation between choices of the equal option and response times reverses as the benefits of the fair option increase: in the VERY LOW game, 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 (Rank-sum test, $p < 0.1$). Hence we classify this game as type 2. In the LOW and HIGH games, the median response times of subjects who chose the equal option are smaller than the response times of subjects who chose the selfish option (Rank-sum test, $p < 0.1$) which is why we classify these games as type 1. There is no significant

¹⁴ 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.

difference in response times for the MEDIUM game (Rank-sum test, $p = 0.64$) which thus constitutes a type 0 game.

We use observed choice frequencies to determine if the DDM makes unambiguous predictions concerning the effect of time pressure in the different games. For the only type 2 situation (VERY LOW), the share of selfish decision is much larger than 50% such that the DDM makes ambiguous predictions regarding the expected effect of time pressure. For the two type 1 situations, the DDM makes unambiguous predictions for the LOW game ($< 50\%$ fair choices) but not for the HIGH game in which a majority of subjects chose the fair option. For the MEDIUM game, the DDM unambiguously predicts that time pressure should not have any effect on the fraction of fair choices given that subjects choose the fair and the selfish option at roughly equal rates (Binomial test, $p = 0.9$). Thus, solely the LOW and the MEDIUM game allow for unbiased tests of the FII hypothesis.

The distribution of response times for the three prisoner's dilemma games are displayed in the bottom panel of Fig. 1.¹⁵ Most importantly, the fraction of cooperators increases with the benefits: only 34% of subjects chose to cooperate in LOW, while this frequency rises to 55% in the MEDIUM and to 58% in the HIGH game. Looking at response times, we find that the median response time of subjects who chose to cooperate is significantly smaller than the median response time of those subjects who chose to defect in each of the three games (Rank-sum test, $p < 0.05$). Thus, the Unconstrained condition only includes PDs of type 1. Looking at the fraction of fair choices, the DDM only makes an unambiguous prediction in the LOW game since the share of cooperators is smaller than 50% in this game. For the MEDIUM and HIGH games, on the other hand, the DDM predictions are not unambiguous and hence they cannot provide unambiguous evidence in favor of or against the FII hypothesis. To also analyze the effect of time pressure in a game that is more likely to be of type 2, we added an additional prisoner's dilemma game (VERY LOW) in the Time Pressure and Time Delay conditions in which we further reduced the benefits of cooperation.¹⁶

4.2 Constrained response time in the binary dictator games

We begin our discussion of the constrained decision time treatments with a series of manipulation checks. Most importantly, time pressure significantly speeds-up decisions across all BD games from an average of 13.34 s (CI: 11.58, 15.11) in the TD to 3.22 s (CI: 3.02, 3.44) in the TP and 2.16 s (CI: 1.97, 2.36) in the STP condition.¹⁷ In addition, average decision times are significantly smaller in the STP as compared to the TP condition (Rank-sum test, $p \leq 0.001$). Game-wise

¹⁵ We later added the VERY LOW game in the subsequent Time Pressure and Time Delay sessions because each of these games represent a type 1 situation.

¹⁶ Since we do not observe response time correlations or choice frequencies for this game, all time pressure results can only be interpreted under the assumption that it indeed represents a type 2 situation in which there is a majority of fair decision makers. The second assumption is unlikely to hold, given that in the LOW game already only 34% of subjects cooperate despite stronger incentives to cooperate.

¹⁷ Note, that the mean decision time in the TD condition is significantly higher than 6 s and only a minority of decisions (13.8%) is made within the range of 6–7 s. This indicates that there are only few

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

	Time delay		Time pressure		<i>p</i> value	Strong time Fisher's		<i>p</i> value
	<i>N</i>	Mean	<i>N</i>	Mean	Exact	<i>N</i>	Mean	Fisher's
VERY LOW	74	0.11	72	0.11	1	30	0.13	0.74
LOW	74	0.32	72	0.44	0.17	30	0.57	0.03
MEDIUM	74	0.53	72	0.60	0.41	30	0.63	0.39
HIGH	74	0.64	72	0.74	0.22	30	0.63	1

The mean rate of fair choices displayed is calculated over all orders. We report the *p* values of a two-sided Fisher's exact test, comparing the fraction of fair choices in the TD condition with the TP (column 6) and the STP condition (column 9)

comparisons furthermore indicate that the effect of time pressure is similar across different games and significantly reduces response times in all four decisions (Rank-sum test, $p \leq 0.001$).¹⁸ Finally, subjects in the TP and STP conditions, who take their first decision under time pressure and their second decision under time delay, are significantly faster (Sign-rank test, $p \leq 0.001$) when taking their first decision as compared to their second decision (TP: 9.66 s [CI: 9.04, 10.30]; STP: 9.51 s [CI: 7.90, 11.13]). Overall, these comparisons indicate that time pressure successfully induced faster decision making among subjects.

In a second manipulation check, we analyze whether subjects indeed perceive the equal option as the fair outcome in all four binary dictator games. For this purpose we analyze subjective (unincentivized) fairness statements elicited at the end of the experiment. We find that in all games, a large majority of subjects perceive the equal option as the fairest outcome (81% in VERY LOW, 97% in LOW, 100% in MEDIUM, 100% in HIGH). We also find that subjective fairness assessments do not differ across the three conditions. This indicates that labeling the equal option as "fair" is strongly in line with the fairness perceptions of our subjects, in particular for the LOW and MEDIUM games that are of most interest for testing the FII hypothesis.

We now turn to analyzing the effect of time pressure on the fraction of fair choices in the BD games. Averaged over all decisions, we do not find evidence that subjects in the time pressure conditions chose the equal "fair" option significantly more often than subjects who took their decision under time delay (40% in TD vs. 47% in TP vs. 49% in STP, Rank-sum test, $p > 0.1$). The main results of our between-subjects test of the FII hypothesis are summarized in Table 5 in which we report the mean fraction of equal "fair" choices in each of the four games separately.

Footnote 17 continued

subjects in the TD condition, who have already completed their decision process when reaching the delay cutoff.

¹⁸ More detailed statistics on response time distributions for each game are available in Online Appendix C.

In the LOW game, the equal “fair” option is not chosen significantly more often when comparing the TD and TP conditions (Two-sided Fisher’s exact test, $p = 0.17$). We obtain a different result, if we restrict our sample to those subjects who took their very first choice in this game (Two-sided Fisher’s exact test, $p = 0.05$). When time pressure is stronger, we do find that the equal allocation is chosen significantly more often in the STP compared to the TD condition (Two-sided Fisher’s exact test, $p = 0.028$), yet this significance vanishes when we restrict our analysis to first choices only (Two-sided Fisher’s exact test, $p = 0.24$). Based on these results we can neither accept nor reject the hypothesis that “fairness is intuitive”. In both treatments we observe that time pressure increases the fraction of fair choices either in the full sample or when restricting our analysis to first choices. Given that in the LOW game both the FII hypothesis and the DDM predict that time pressure should increase the fraction of fair choices, the observed increases in fair choices can, however, at most provide ambiguous evidence in favor of the FII hypothesis.

According to the DDM, time pressure should have no effect on the fraction of fair choices in the MEDIUM game. Hence an increase in fairness would be unambiguous evidence in favor of the FII. The fraction of subjects who select the equal option in the TP and STP conditions is indeed slightly higher under time pressure as compared to the TD condition, but this increase is not significant (Two-sided Fisher’s exact test; TP: $p = 0.41$, STP: $p = 0.39$). Restricting our comparison to first choices only does not alter this result (Two-sided Fisher’s exact test; TP: $p = 0.76$, STP: $p = 0.10$). These results constitute unambiguous evidence against the FII hypothesis.

In the VERY LOW and HIGH game we find no evidence that time pressure affects the frequency of fair choices (Two-sided Fisher’s exact test, $p > 0.1$). Since for these games it is unclear if the DDM predicts effects that are in line or orthogonal to the FII, we cannot unambiguously reject the FII based on these observations.

Result 1 (Between-Subject evidence in BDs) *We find evidence that time pressure increases the fraction of fair choices in games, in which this increase can be accounted for by both the DDM and the FII hypothesis (LOW game). In contrast, we do not find that time pressure significantly increases the fraction of fair choices in games, in which such increase can only be accounted for by the FII hypothesis (MEDIUM game).*

Our design also allows assessing within-subjects evidence by comparing a subject’s initial choice under time pressure and her second choice (in the same game) under time delay. The two games that can provide unambiguous evidence in favor of or against the FII hypothesis are the LOW and the MEDIUM game. 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 in the Time Pressure conditions to the switching rates in the Time Delay condition, where subjects take both decisions under time delay. To analyze switching behavior, we computed a variable that takes a value of 1 if a subject switched from the fair (first choice) to the selfish option

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

	Time delay			Time pressure			Rank-sum <i>p</i> value	Strong Time Pressure			Rank-sum <i>p</i> value
	<i>N</i>	Mean	Sign-rank	<i>N</i>	Mean	Sign-rank		<i>N</i>	Mean	Sign-rank	
	<i>(within-subjects)</i>			<i>(within-subjects)</i>			<i>(across)</i>	<i>(within-subjects)</i>			<i>(across)</i>
VLOW	74	0.04	0.18	72	0.03	0.41	0.78	30	0.1	0.08	0.32
LOW	74	0.07	0.19	72	0.15	0.02	0.27	30	0.27	0.01	0.05
MEDIUM	74	0.03	0.59	72	-0.03	0.64	0.48	30	0	1	0.80
HIGH	74	-0.01	0.74	72	0.11	0.03	0.05	30	-0.03	0.71	0.81

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 and STP condition and under time delay in the TD condition. The decision in block 2 is taken under time delay in all three conditions. Columns 4,7 and 11 report the *p* value of a Wilcoxon Sign Rank Test, performed on subject's switching behavior within one condition. In addition, we report the *p* value of a Rank-sum test, which compares switching behavior across the TD and TP (column 8) or STP (column 12) conditions

(second choice), and a value of -1 if a subject switched from the selfish to the fair option (see Table 6).

First, looking at the switching rates in the two time pressure conditions, we do find that subjects are 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 LOW, but not in the MEDIUM game. The former finding is consistent with the DDM and the FII hypothesis while 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 caused by our treatment, we compare the switching rate in the LOW and MEDIUM games in the two Time Pressure conditions to the Time Delay condition. The results of the Rank-sum test are reported in Table 6. Our analysis shows that in the LOW game, subjects in the STP condition were indeed more likely to switch from the equal to the unequal option compared to subjects in the TD condition. However, given that the DDM and the FII hypothesis both support this prediction, the evidence can only provide ambiguous evidence in favor of the FII hypothesis. In contrast, there is no statistical significant difference in switching patterns for subjects in the TP and TD conditions. This result constitutes unambiguous evidence against the FII hypothesis. For the MEDIUM game we find no evidence that there is significantly more switching behavior under time pressure than under time delay. This is direct evidence against the FII hypothesis.¹⁹

The interpretation of the previous results rests on the assumption that we indeed classified the games correctly. As a robustness check, we employ a complementary within-subject test that does not depend on the classification of the games but

¹⁹ We do not find evidence that there are differences in switching behavior for the VERY LOW game, while there is significantly more switching for the HIGH game in the TP but not in the STP condition. Since both of these game can only provide ambiguous evidence in favor or against the FII hypothesis, we are not discussing these results in more detail.

instead exploits the fact that we observe choices in four different games per subject. Based on these four choices we infer in which game a subject should be closest to her individual indifference point.²⁰ Let $C_i = (x_1, x_2, x_3, x_4)$ describe the set of choices that a subject makes in the first four games such that $C_0 = (F, F, F, F)$ describes a subject who chooses the fair option in all four games and $C_2 = (S, S, F, F)$ describes a subject who chooses the selfish option only in the VERY LOW and LOW game. A C_0 subject is closest to her indifference point in the VERY LOW game since in this game choosing the fair option is more costly than in any of the other games. A C_2 subject is closest to her indifference point in the LOW or MEDIUM game since she switches from the selfish to the fair option between these two games. Overall, there are five different choice patterns that allow to approximate the location of the indifference point and 84.66% of subjects can be classified according to these patterns.²¹ This classification can shed light on the role of subjective choice difficulty in the following way: the DDM predicts that subjects are more likely to make a mistake in games in which they are closer to indifference. Thus, when comparing a subject's first choices with her second choices in the same four games, she should be more likely to switch options in games closer to her indifference point. The FII, on the other hand, predicts that subjects should display a similar rate of switching for all games in which they have initially selected the fair option. Furthermore, there should be few to no switches in games in which subjects have initially selected the selfish option. These two predictions can be easily illustrated using an exemplary subject: assume a subject has chosen $C_1 = (S, F, F, F)$ in the initial four games and is classified accordingly. The FII hypothesis predicts that when comparing the subject's first to her second choices in the same games, she should switch to the selfish option with the same probability in the LOW, MEDIUM and HIGH games. Conversely, according to the DDM, the highest frequency of switches should occur in the VERY LOW or LOW game—as the exemplary subject is closest to her indifference point in these games—while there should be fewer switches in the MEDIUM and HIGH games. Figure 2 displays the propensity to switch in each game for each choice pattern.²²

For almost all classifications, switching patterns are more closely in line with predictions of the DDM than with predictions of the FII hypothesis. With the exception of the *SSFF* pattern, we observe that subjects are more likely to switch in games which are closer to their indifference point. In contrast, we find little evidence that subjects are switching at similar rates in all games in which they have chosen the fair option under time pressure. This is most evident for the *FFFF*

²⁰ This method rests on the assumption that subjects have transitive preferences over own-other allocations that are only violated by mistake (Andreoni and Miller 2002).

²¹ The five choice patterns are (F, F, F, F), (S, F, F, F), (S, S, F, F), (S, S, S, F) and (S, S, S, S). For any other pattern [e.g. (S, F, S, F)] there is no clear indication in which decision a subject might have made an error that violates transitivity and hence where the indifference point for this subject might be located. Note that a consistent pattern does not necessarily imply that subjects have not made any error since a (S, F, F, F) subject could have made an error in either the VERY LOW game implying that her actual preferences are (F, F, F, F) or in the LOW game implying that her actual preferences are (S, S, F, F) or could have made more than one error.

²² Due to the smaller group size for some classifications in the STP condition, we pooled data from both time pressure conditions for this analysis.

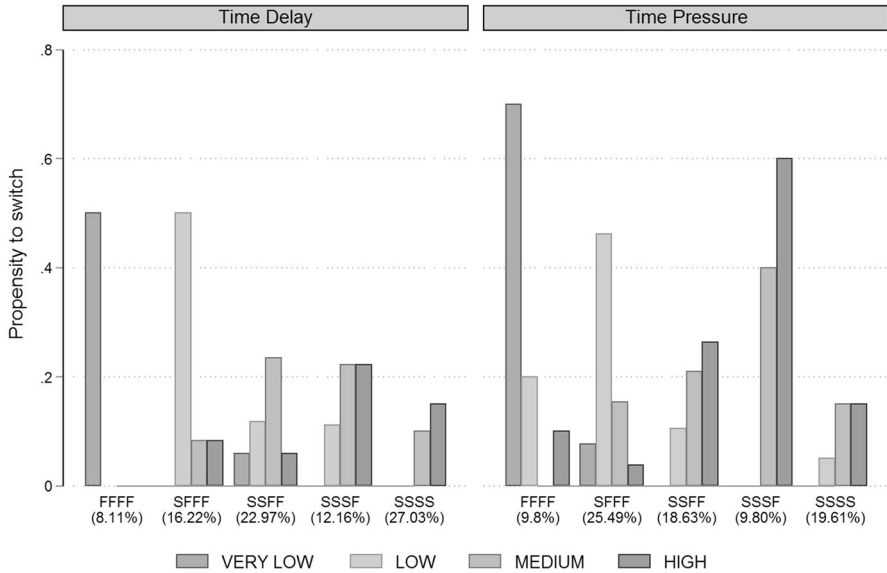


Fig. 2 Conditional switching probabilities in the BDs.

Note: This figure displays the propensity to switch between first and second choice in a given game across five different classifications of consistent decision making. The left panel displays switching behaviour when first choices have been made under time delay and the right panel displays switching behaviour when first choices have been made under time pressure. The percentages indicate how common a classification is within the population

pattern, where a majority of switches occur in the first (VERY LOW) game even under time pressure. In contrast to the predictions of the FII hypothesis, there is also substantive evidence for switching from the selfish to the fair option (most pronounced for the SSSF and SSSS choice patterns). Given that most of these switches occur for games that are close to individual indifference points, this pattern is closely in line with the predictions of the DDM.

Result 2 (Within-Subject evidence in the BDs) *We do find evidence that subjects are more likely to switch from the fair to the selfish option in games, in which this prediction is supported by the FII and the DDM (LOW); but not in games, in which this prediction is only supported by the FII hypothesis (MEDIUM). The latter provides unambiguous evidence against 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 consistent evidence in favor of the FII hypothesis using only the first decision taken by each subject. 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. We report the results of three different specifications in Table 7. In specification (1), we find no evidence that time pressure increases the frequency of equal choices, when controlling for order effects and the benefits of choosing the fair option. In specification (2), we add interaction terms between the treatment dummy and the benefits of choosing the fair option. Again, we find no evidence that time pressure significantly affects equal choices in any of the four games. Moreover, all interaction terms for the standard time pressure (TD) treatment are insignificant. This is further evidence that even in those games where both the DDM and the FII would predict an increase of fairness under time pressure there is no such effect. For stronger time pressure (STP), there is weakly significant evidence that time pressure increases the frequency of fair choices in the LOW game, but not in the MEDIUM game. The former finding is, however, predicted by both the DDM and the FII hypothesis. 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. As suspected, we do observe that the screen variables are significant in all three specifications. Thus, if a game was presented on a specific decision screen, the likelihood that a subject would choose the equal option was increased or decreased depending on the specification.

4.3 Constrained response time in the prisoner's dilemma games

As for the BD games, we find that time pressure significantly speeds up choices in the PD games.²³ Furthermore subjects' individual fairness assessments, which we elicited at the end of the experiment, are strongly consistent with our label: a large majority of the subjects perceives the cooperative option as the fairest outcome in all four prisoner's dilemma games (VLOW 94%, LOW 96%, MEDIUM 88%, HIGH 97%), independent of the treatment condition.

Based on the analysis of response times and choice frequencies in the Unconstrained condition, the only PD game that can shed light on the FII hypothesis is the LOW game. The fraction of cooperative "fair" choices in the LOW game increases from 41% in the TD condition to 44% in the TP and 50% in the STP condition. This increase is not statistically significant though (Two-sided Fisher's exact test; TD vs. TP: $p = 0.74$, TD vs. STP: $p = 0.39$). When we restrict our analysis to choices on the first decision screen (TD: 69% vs. TP: 60% vs. STP: 50%), we again find no evidence that time pressure increases the fraction of fair behavior but rather observe a slight decrease (Two-sided Fisher's exact test; TD vs. TP: $p = 0.73$, TD vs. STP: $p = 0.43$). Given that in the LOW game the FII and the DDM both predict that time pressure should increase the fraction of fair choices, this observation is unambiguous evidence against the hypothesis that "fairness is intuitive" and the findings in Rand et al. (2012).

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

²³ All response time statistics are available in Online Appendix C.

Table 7 Random effects probit regression for the effect of time pressure in the binary dictator games

	(1)	(2)	(3)
Time pressure (TP)	0.350 (1.33)	0.00634 (0.01)	0.249 (0.43)
Strong Time Pressure (STP)	0.463 (1.30)	0.115 (0.20)	0.261 (0.33)
LOW	1.986**** (7.11)	1.632**** (4.28)	1.707**** (4.20)
MEDIUM	2.574**** (9.18)	2.480**** (6.26)	2.524**** (6.22)
HIGH	2.773**** (9.40)	2.627**** (6.28)	2.698**** (5.93)
SCREEN2	- 0.345* (- 1.80)	- 0.382* (- 1.95)	- 0.259 (- 0.87)
SCREEN3	0.251 (1.41)	0.259 (1.44)	0.622** (2.02)
SCREEN4	0.176 (0.71)	0.197 (0.77)	0.239 (0.63)
TP * LOW		0.541 (1.12)	0.392 (0.69)
TP * MEDIUM		0.242 (0.49)	0.176 (0.33)
TP * HIGH		0.489 (0.88)	0.436 (0.77)
STP * LOW		1.010* (1.78)	1.273 (1.60)
STP * MEDIUM		0.291 (0.49)	0.567 (0.83)
STP * HIGH		- 0.0806 (- 0.13)	- 0.0736 (- 0.11)
TP * SCREEN2			- 0.0606 (- 0.14)
TP * SCREEN3			- 0.621 (- 1.51)
TP * SCREEN4			- 0.0298 (- 0.05)
STP * SCREEN2			- 0.602 (- 1.04)
STP * SCREEN3			- 0.598 (- 1.26)
STP * SCREEN4			0.0114 (0.02)

Table 7 continued

	(1)	(2)	(3)
CONSTANT	– 2.396**** (– 7.35)	– 2.231**** (– 5.85)	– 2.410**** (– 5.57)
Observations	704	704	704
Subjects	176	176	176
Prob > χ^2	0.0000	0.0000	0.0000

The dependent variable takes a value of 1 if the equal option was chosen and 0 otherwise. The variables Time Pressure and Strong Time Pressure equal 1 if the subject was assigned to the corresponding condition and 0 otherwise. Low, Medium and High are dummy variables for the games with which we confronted subjects in block 1 and effects are reported 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 effects are reported relative to the decision on the first screen

t statistics in parentheses

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

Time Delay compared to the Time Pressure or Strong 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. We find that in the LOW game, average beliefs did not differ between the two conditions (Rank-sum test; TD vs. TP: $p = 0.63$, TD vs. STP: $p = 1$).

In a second step, we analyze the within-subject effect of time pressure on choices in the LOW prisoner's dilemma game. For this purpose, we compute switching probabilities by comparing a subject's first choice with her second choice in the same game. 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. Note that the DDM makes the same prediction. Thus, if we do not observe the expected switching pattern, this would constitute unambiguous evidence against the FII hypothesis.

The first thing to note is that subjects in the TP and STP conditions are indeed more likely to switch from cooperation under time pressure to defection under time delay (Sign Rank Test; TP $p = 0.07$, STP $p = 0.01$). Subjects in the TD condition are also more likely to switch from cooperation to defection, but this difference is not significant (Sign Rank Test; $p = 0.2$). When we compare switching behavior in each of the two Time Pressure to switching behavior in the Time Delay condition, we do not find that subjects in either of the two time pressure conditions were more likely to switch from cooperation to defection as compared to subjects in the Time Delay condition (Rank-sum test; TD vs. TP: $p = 0.62$, TD vs. STP: $p = 0.12$). Hence, instead of reflecting a reassessment of an initial intuitive decision, the decline of cooperative choices in the Time Pressure conditions might simply reflect the well-known fact that subjects tend to become more selfish in repeated decisions even without receiving feedback (Ledyard Unpublished). Therefore, our within-subject evidence in in this game does not support the FII hypothesis.

Like in the BD games, a complementary analysis of conditional switching patterns at the individual level shows that most switches occur in games in which subjects are close to their indifference point. These observations support the idea that switching behavior under time pressure reflects choice difficulty instead of a reassessment of an intuitive fair choice.²⁴

Result 3 (Between- and within-subject evidence in the PDs) *In the Prisoner's dilemma games, we do not find evidence that time pressure increases the fraction of fair choices even when both the FII as well as the DDM would support this prediction (LOW game). Within-subjects, we find that subjects in both Time Pressure conditions are as likely to revise an initially fair choice as subjects in the Time Delay condition. Both results are inconsistent with the FII hypothesis.*

5 Conclusion and discussion

In this paper we propose and conduct a new test of the FII hypothesis (Rand et al. 2012; Cappelen et al. 2016). Our test takes into account that a causal test of this hypothesis, using time pressure and time delay manipulations, 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 frequency 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 and depending on the distribution of preference types within the population. 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 BD or PD games classified accordingly, we find that time pressure consistently increases the fraction of fair choices. On the other hand, we do not find that time pressure increases the fraction of fair choices in games, in which this increase is only predicted by the FII hypothesis, thus rendering unambiguous evidence against the FII hypothesis. Our empirical test therefore provides little support for the hypothesis that “fairness is intuitive” in a general way. This result holds between- and within-subjects. A complementary analysis further demonstrates that switching patterns strongly reflect choice difficulty (subjective indifference), a pattern that is supported by the DDM but not by the FFI hypothesis.

On the one hand our rejection of the FII hypothesis is in line with a number of recent papers (Fiedler et al. 2013; Tinghög et al. 2013, 2016; Martinsson et al. 2014; Duffy and Smith 2014; Verkoeijen and Bouwmeester 2014; Achtziger et al. 2015; Kocher et al. 2016; Lohse 2016; Capraro and Cococcioni 2016) and a large scale replication project (Bouwmeester et al. 2017) which also suggest that in some instances behaving fairly might not be intuitive and might even require additional

²⁴ The full analysis of the remaining games and switching patterns can be found in Online Appendix C.

deliberation or stronger self-control. 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. 2016). 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).²⁵ The observation that some experiments have found an increase of fairness under time pressure while other experiments report no effect or even a reduction of fairness 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 conceivable 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 analyze 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 possible that subjects with different individual attributes or cultural backgrounds might attach different subjective valuations to options in the same task, thereby leading to unobserved heterogeneity in terms of the perceived choice difficulty as well as the share of fair decision makers. Given that both of these factors determine if the DDM predicts an increase or decrease of fair behavior under time pressure, these experiments might come to different conclusions concerning the FII hypothesis.

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, default options) 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,²⁶ 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 and consequently substantially higher. We believe that each of

²⁵ For a discussion of the last paper see Myrseth and Wollbrant (2016) and Jagau and van Veelen (2017).

²⁶ 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.

these 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 in an obvious way if it was generally valid, as suggested by the mechanism motivating 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 further moderating factors might provide useful insights into the conditions or individual attributes that influence the link between intuition and fairness. Several recent contributions have already provided first insights into the role of confusion (Recalde et al. 2014; Stromland et al. 2016; Goeschl and Lohse 2016), gender (Rand et al. 2016; Tinghög et al. 2016), culture (Nishi et al. 2017), stake size (Mrkva 2017) and social-value-orientation (Chen and Fischbacher 2015; Mischkowski and Glöckner 2016).

Acknowledgements We would like to thank Christoph Vanberg, Gustav Tinghög, Ariel Rubinstein, Ernst Fehr, Tobias Pfommer, Daniel Heyen, Gert Pönitzsch, 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. In addition, we would like to thank the editors as well as two anonymous referees for their tremendously helpful feedback on an earlier version of this paper.

References

- Achtziger, A., Alós-Ferrer, C., & Wagner, A. K. (2015). Money, depletion, and prosociality in the dictator game. *Journal of Neuroscience, Psychology, and Economics*, 8(1), 1.
- Alós-Ferrer, C. (2016). A dual-process diffusion model. *Journal of Behavioral Decision Making*. <https://doi.org/10.1002/bdm.1960>.
- Alós-Ferrer, C., & Strack, F. (2014). From dual processes to multiple selves: Implications for economic behavior. *Journal of Economic Psychology*, 41, 1–11.
- Andreoni, J., & Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2), 737–753.
- Bear, A., & Rand, D. G. (2016). Intuition, deliberation, and the evolution of cooperation. *Proceedings of the National Academy of Sciences*, 113(4), 936–941.
- Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, 71, 117–120.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), 700.
- Bouwmeester, S., Verkoeijen, P. P., Aczel, B., Barbosa, F., Bègue, L., Brañas-Garza, P., et al. (2017). Registered replication report: Rand, greene, and nowak (2012). *Perspectives on Psychological Science*, 12(3), 527–542.
- Buckert, M., Oechssler, J., & Schwieren, C. (2017). Imitation under stress. *Journal of Economic Behavior and Organization*, 139, 252–266.
- Caplin, A., & Martin, D. (2015). The dual-process drift diffusion model: Evidence from response times. *Economic Inquiry*, 54(2), 1274–1282.
- Cappelen, A. W., Nielsen, U. H., Tungodden, B., Tyran, J.-R., & Wengström, E. (2016). Fairness is intuitive. *Experimental Economics*, 19, 727–740.
- Capraro, V., & Cococcioni, G. (2016). Rethinking spontaneous giving: Extreme time pressure and ego-depletion favor self-regarding reactions. *Scientific Reports*, 6, 27219.

- Capraro, V., Corgnet, B., Espín, A. M., & Hernán-González, R. (2017). Deliberation favours social efficiency by making people disregard their relative shares: Evidence from USA and India. *Royal Society Open Science*, 4(2), 160605.
- Chen, F., & Fischbacher, U. (2015). Cognitive processes of distributional preferences: A response time study. *Research Paper Series Thurgauer Wirtschaftsinstitut*.
- Clithero, J. A. (2016). Response times in economics: Looking through the lens of sequential sampling models. Available at SSRN. <https://doi.org/10.2139/ssrn.2795871>.
- Cubitt, R. P., Drouvelis, M., Gächter, S., & Kabalin, R. (2011). Moral judgments in social dilemmas: How bad is free riding? *Journal of Public Economics*, 95(3), 253–264.
- Dreber, A., Fudenberg, D., Levine, D. K., & Rand, D. G. (2014). Self-control, social preferences and the effect of delayed payments. Available at SSRN. <https://doi.org/10.2139/ssrn.1752366>.
- Drouvelis, M., & Grosskopf, B. (2016). The effects of induced emotions on pro-social behaviour. *Journal of Public Economics*, 134, 1–8.
- Duffy, S., & Smith, J. (2014). Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *Journal of Behavioral and Experimental Economics*, 51, 47–56.
- Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *The American Economic Review*, 94(4), 857–869.
- Faravelli, M. (2007). How context matters: A survey based experiment on distributive justice. *Journal of Public Economics*, 91(7), 1399–1422.
- Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism experimental evidence and new theories. *Handbook of the Economics of Giving, Altruism and Reciprocity*, 1, 615–691.
- Fiedler, S., Glöckner, A., Nicklisch, A., & Dickert, S. (2013). Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational Behavior and Human Decision Processes*, 120(2), 272–284.
- Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178.
- Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19(4), 25–42.
- Goeschl, T., & Lohse, J. (2016). Cooperation in public good games. Calculated or confused? *AWI Discussion Paper Series No 626*.
- Hawkins, G. E., Forstmann, B. U., Wagenmakers, E.-J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience*, 35(6), 2476–2484.
- Hieber, P., & Scherer, M. (2012). A note on first-passage times of continuously time-changed brownian motion. *Statistics & Probability Letters*, 82(1), 165–172.
- Hopfensitz, A., & Reuben, E. (2009). The Importance of Emotions for the Effectiveness of Social Punishment. *The Economic Journal*, 119(540), 1534–1559.
- Jagau, S., & van Veen, M. (2017). A general evolutionary framework for the role of intuition and deliberation in cooperation. *Nature Human Behavior*, 1, 0152.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697.
- Kocher, M. G., Martinsson, P., Myrseth, K. O. R., & Wollbrant, C. E. (2016). Strong, bold, and kind: Self-control and cooperation in social dilemmas. *Experimental Economics*, 20(1), 44–69.
- Krajbich, I., Bartling, B., Hare, T., & Fehr, E. (2015a). Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications*, 6, 7455.
- Krajbich, I., Hare, T., Bartling, B., Morishima, Y., & Fehr, E. (2015b). A common mechanism underlying food choice and social decisions. *PLoS Computational Biology*, 11(10), e1004371.
- Krajbich, I., Oud, B., & Fehr, E. (2014). Benefits of neuroeconomic modeling: New policy interventions and predictors of preference. *The American Economic Review*, 104(5), 501–506.
- Ledyard, J. (Unpublished). Public goods: A survey of experimental research. Social Science Working Paper, 861. California Institute of Technology, Pasadena, CA.
- Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *The American Economic Review*, 90(2), 426–432.
- Lohse, J. (2016). Smart or selfish—When smart guys finish nice. *Journal of Behavioral and Experimental Economics*, 64(10), 28–40.
- Martinsson, P., Myrseth, K. O. R., & Wollbrant, C. (2014). Social dilemmas: When self-control benefits cooperation. *Journal of Economic Psychology*, 45, 213–236.

- Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The drift diffusion model can account for value-based choice response times under high and low time pressure. *Judgment and Decision Making*, 5(6), 437–449.
- Mischkowski, D., & Glöckner, A. (2016). Spontaneous cooperation for prosocials, but not for proselves: Social value orientation moderates spontaneous cooperation behavior. *Scientific Reports*, 6, 21555.
- Mrkva, K. (2017). Giving, fast and slow: Reflection increases costly (but not uncostly) charitable giving. *Journal of Behavioral Decision Making*, 30(5), 1052–1065.
- Myrseth, K. O. R., & Wollbrant, C. E. (2016). Models inconsistent with altruism cannot explain the evolution of human cooperation. *Proceedings of the National Academy of Sciences*, 113(18), E2472.
- Myrseth, K. O. R., & Wollbrant, C. E. (2017). Cognitive foundations of cooperation revisited: Commentary on rand et al. (2012, 2014). *Journal of Behavioral and Experimental Economics*, 69, 133–138.
- Nishi, A., Christakis, N. A., & Rand, D. G. (2017). Cooperation, decision time, and culture: Online experiments with american and indian participants. *PLoS ONE*, 12(2), e0171252.
- Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, 5(5), 1–1.
- Polanía, R., Krajbich, I., Grueschow, M., & Ruff, C. C. (2014). Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron*, 82(3), 709–720.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review*, 83(5), 1281–1302.
- Rand, D. G., Brescoll, V. L., Everett, J. A., Capraro, V., & Barcelo, H. (2016). Social heuristics and social roles: Intuition favors altruism for women but not for men. *Journal of Experimental Psychology: General*, 145(4), 389.
- Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489(7416), 427–430.
- Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., et al. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, 5, 3677.
- Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling accumulation of partial information. *American Psychological Association*, 95(2), 238–255.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9(5), 347–356.
- Recalde, M. P., Riedl, A., & Vesterlund, L. (2014). Error prone inference from response time: The case of intuitive generosity. *CESifo Working Paper Series No. 4987*.
- Reuben, E., & Riedl, A. (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, 77(1), 122–137.
- Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *The Economic Journal*, 117(523), 1243–1259.
- Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology*, 44(3), 408–463.
- Spiliopoulos, L., & Ortmann, A. (2017). The BCD of response time analysis in experimental economics. *Experimental Economics*. <https://doi.org/10.1007/s10683-017-9528-1>.
- Stromland, E., Tjotta, S., & Torsvik, G. (2016). Cooperating, fast and slow: Testing the social heuristics hypothesis. *CESifo Working Paper Series No. 5875*.
- Tinghög, G., Andersson, D., Bonn, C., Böttiger, H., Josephson, C., Lundgren, G., et al. (2013). Intuition and cooperation reconsidered. *Nature*, 498(7452), 427–430.
- Tinghög, G., Andersson, D., Bonn, C., Johannesson, M., Kirchler, M., Koppel, L., et al. (2016). Intuition and moral decision-making—the effect of time pressure and cognitive load on moral judgment and altruistic behavior. *PLoS ONE*, 11(10), e0164012.
- Verkoeijen, P. P., & Bouwmeester, S. (2014). Does intuition cause cooperation? *PLoS ONE*, 9(5), e96654.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, 32(7), 1206–1220.
- Wright, P. (1974). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology*, 59(5), 555–561.