# The Role of Attention and Frames on Third-Party Punishment and Compensation Choices

Claudia Civai, Valerio Capraro, and Luca Polonio

## Supplementary Materials

## 1. Experiment one

### 1.1 Eye-tracking measurement

To measure eye movement, we used desktop mount Eyelink 1000 (SR Research, Canada) with a sampling rate of 1000 Hz. A fixation was defined as an interval in which gaze was focused within 1 degree of visual angle for at least 60 ms. Participants rested their chin and forehead on a chinrest at approximately 60 cm from a monitor with 1920 X 1980 resolution. The calibration and validation procedure[1] of the eye-tracking was repeated at the beginning of each block (we had two blocks in total, although participants could ask to rest after any trial, without moving their head). We used a 13-point calibration. The calibration phase was repeated until the difference between the positions of the points on the screen and the corresponding eye locations was less than 1°. After the calibration phase, a validation phase was executed to make sure that the calibration had been accurate. Recalibration and revalidation were performed if these had been unsuccessful.

### 1.2 Descriptive statistics.

*Table 1. Experiment one. Descriptive statistics for the variables in the lab-based eye-tracking experiment. Punish and Compens variables refer to the proportion of punishment and compensation choices across all trials (including the "Leave" trials); Amount refers to the*

[1] A standard two-step procedure that allows the eye-tracker system to track the gaze location on the computer screen.

*average amount (chips) spent across trials where either punishment or compensation was chosen; A_B_Dwell refers to the difference between the time spent looking at A and B.*

| | Valid | Missing | M | SD | Min | Max |
|---|---|---|---|---|---|---|
| Punish_coin | 36 | 0 | 0.378 | 0.19 | 0.03 | 0.84 |
| Punish_digit | 36 | 0 | 0.366 | 0.21 | 0.02 | 0.91 |
| Compens_coin | 36 | 0 | 0.248 | 0.16 | 0 | 0.52 |
| Compens_digit | 36 | 0 | 0.259 | 0.17 | 0 | 0.62 |
| Amount_punish_coin | 36 | 0 | 68.39 | 19.26 | 10 | 100 |
| Amount_punish_digit | 36 | 0 | 65.49 | 17.40 | 20 | 100 |
| Amount_compens_coin | 36 | 0 | 53.04 | 27.16 | 0 | 93.85 |
| Amount_compens_digit | 36 | 0 | 51.14 | 27.52 | 0 | 90.53 |
| A_B_Dwell_coin | 36 | 0 | 0.144 | 0.20 | -0.07 | 0.83 |
| A_B_Dwell_digit | 36 | 0 | 0.052 | 0.1 | -0.11 | 0.35 |

### 1.3 Manipulation check.

The likelihood of punishment and compensation (punishment: est. = 0.73, s.e. = 0.04, $z$ = 20.15, $p < .001$; compensation: est. = 0.62, s.e. = 0.04, $z = 16.56$, $p < .001$), as well as the amount spent to punish and compensate (punish: $\beta = 0.01$, s.e. = 0.01, $t(1691) = 18.53$, $p < .001$; compensate: $\beta = 0.01$, s.e. = 0.01, $t(1149) = 14.83$, $p < .001$), increase with the increase of the level of injustice (Figures 2b and 2c in the main text), showing that people are more likely to react (vs non-react), and spend more to react with the increase of injustice.

### 1.4 Alternative hypothesis 3 testing

An alternative mixed model was run to partially test hypothesis 3, i.e., that the condition influences the amount spent. Specifically, we considered the predictive effects of injustice, difference in dwell time, condition, and response (punishment vs compensation, excluding trials in which participants selected Leave), on the amount spent. Results show that the higher the injustice the more participants would spend, and that participants spent more to punish (vs compensate), and spent more in the coin (vs digit) condition. No effect of difference in dwell

time was found; however, difference in dwell time does not predict the amount spent per se, but the amount spent to punish (or compensate). Figure 1 shows the estimates, including error bars and p-values.

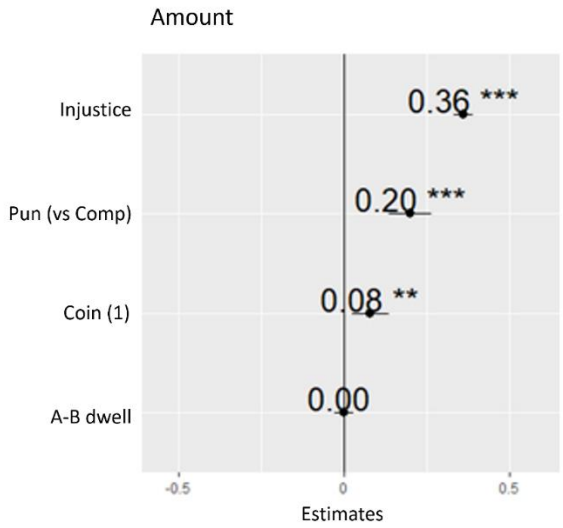


**Figure 1.** Experiment one – lab-based experiment. Standardised estimates of the effects injustice level, Response (punish vs compensate), condition and A-B dwell time on the amount spent; error bars (95% CI) and significance (*$p < .05$; ***$p < .001$) are also reported, with the vertical black line indicating null effect.

### 1.5 Results for the exploratory analysis

A paired samples t-test confirmed the results of the mixed model analysis: the difference in the proportion of punishment and compensation between the coins and the digits condition is not significant ($t(35) = 1.13$, $p = .267$, $d = .19$, 95% CI [-0.02, 0.07]), but the difference between the coins and the digits condition in the amount spent to punish is ($t(35) = 2.45$, $p = .019$, $d = .41$, 95% CI [0.50, 5.29]). We used the effect size of the first t-test to inform the sample size determination of the online experiment (Experiment 2).

## 2. Experiment Two

**2.1 Differences between experiment one and two.** There were some differences between the lab-based and the online experiment, which are reported below.

The number of trials was reduced, considering the increased number of participants and the need to limit the length of the experiment in the online setting. To achieve this, the "fair" trials were removed, and only the unfair trials, i.e., trials in which the offender takes some chips from the victim, were included; moreover, to increase the number of trials without repeating the same outcomes, the levels of injustice were increased, so that instead of 0, 25, 50, 75, 100, we had 25, 50, 75, 100, 125, 150, 175, 200; therefore, each condition (coins/digits) had eight trials, with varying levels of injustice. Overall, 16 trials were presented, and the offender's payoff was presented on the right (or left) four times per condition.

The option "Leave" was removed: participants had to choose between two options, i.e., Take from A (punish) or Give to B (compensate), and their position was counterbalanced between participants. The aim of the study was to understand the relationship between attention and preference for punishment vs compensation, rather than preference to react or not to react to injustice, and therefore we opted to force participants to choose one of the two options of interest. This is also why fair offers were dropped, as it would not be feasible to ask participants to choose between punishment and compensation for fair trials. To offset this forced choice, we added the option to spend 0 to punish or compensate when choosing the amount.

The online experiment was not incentivised: participants were paid a show-up fee via Prolific, but then the scenarios were presented as hypothetical, and therefore there was no deception. Since the aim was to investigate the differences between two ways of reacting to injustice, we considered that self-interest was not necessarily playing a role in the decision.

**2.2 Mixed model analysis**

In line with the lab-based experiment, we employed mixed models where we accounted for the random effect (intercept) of participants:

- Model 2.1: a linear mixed model was run to test whether the condition (coins vs digits) predicted the difference between the percentage of dwell time on the offender's and the victim's payoffs (Hypothesis 1). The full results for this model are reported in Figure 2a.
- Model 2.2: a logistic mixed model was run to test whether the difference between the percentage of dwell time on the offender's and the victim's payoffs (Hypothesis 2) and the condition (coins vs digits) (Hypothesis 3) predict the likelihood to punish vs compensate. The full results for this model are reported in Figure 2b.
- Models 2.3: Two additional linear mixed models were also run to test for the predictive effects of the difference between the dwell time on the offender's and the victim's payoffs and the condition on the amount spent to punish and compensate. The full results for these models are reported in Figure 2c.

**2.3 Descriptive statistics.**

*Table 2. Experiment two. Descriptive statistics for the variables in the online eye-tracking experiment. Punish variable refers to the proportion of punishment choices; Amount refers to the average amount spent to punish or compensate; A_B_Dwell refers to the difference between the time spent looking at A and B.*

| | Valid | Missing | M | SD | Min | Max |
|---|---|---|---|---|---|---|
| Punish_coin | 201 | 0 | 0.67 | 0.24 | 0 | 1 |
| Punish_digit | 201 | 0 | 0.65 | 0.25 | 0 | 1 |
| Amount_punish_coin | 201 | 0 | 55.61 | 25.59 | 0 | 100 |
| Amount_punish_digit | 201 | 0 | 53.23 | 25.84 | 0 | 100 |
| Amount_compens_coin | 201 | 0 | 38.42 | 28.62 | 0 | 100 |
| Amount_compens_digit | 201 | 0 | 40.96 | 29.05 | 0 | 100 |
| A_B_Dwell_coin | 201 | 0 | 2.59 | 5.65 | -11.75 | 27.37 |
| A_B_Dwell_digit | 201 | 0 | 1.57 | 5.71 | -15.375 | 19.75 |

**2.4 Manipulation check.**

As expected, the amount spent to punish ($\beta$ = 0.25; s.e. = 0.01, $t(3011) = 17.6$, $p < .001$) and compensate ($\beta$ = 0.19; s.e. = 0.01, $t(3011) = 12.97$, $p < .0014$) is affected by the level of injustice, or amount of chips taken from B (Figure 2c below), showing that the more severe the injustice is, the more people spend to punish or compensate.

**2.5 Mixed model results.**

*Hypothesis 1: The difference in the duration of gaze directed towards the offender versus the victim is larger in the coin condition compared to the digits condition.* Model 2.1 showed that the difference between the duration of gaze towards the offender and the victim was marginally predicted by the condition ($\beta$ = 0.06; s.e. = 0.03, $t(3013) = 1.85$, $p = .064$) (Figure 2a).

*Hypothesis 2: The more participants look at the offender's payoff, as opposed to the victim's, the more likely they are to punish and the more they spend on punishment.* As expected, model 2.2 shows that the likelihood of punishment increases with the increase of the difference between duration of gaze towards the offender A versus victim B (55% more likely to punish, hence more than 50% chance level, when participant looks at the offender 1% more than the victim; est. = 0.26, s.e. = 0.04, $z = 4.73$, $p < .001$) (Figure 2b). Note that since in this case only two options were available, punishment and compensation were perfectly collinear.

*Hypothesis 3: Participants are more likely to punish, and spend more to punish, in the coins condition compared to the digits condition.* As for the lab-based experiment, model 2.2 showed no effect of condition on the choice to punish (est. = 0.09, s.e. = 0.08, $z = 1.179$, $p = .238$), whilst models 2.3 showed that the participants spent more to punish in the coins condition compared to the digits condition ($\beta$ = 0.07, s.e. = 0.03, $t(3011) = 2.567$, $p = .01$),

but no effect was found on compensation amount ($\beta$ = -0.02, s.e. = 0.03, $t$(3011) = -0.82, $p$ = .412) (Figure 2c).

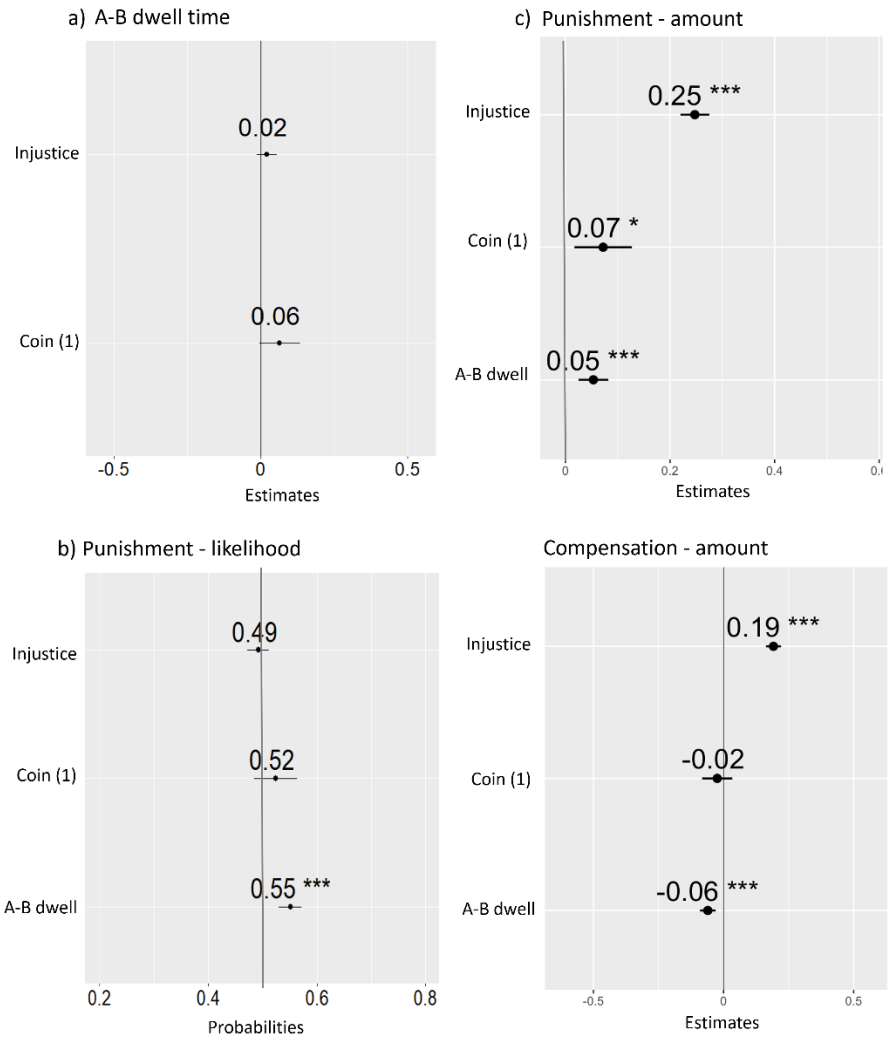


**Figure 2.** Experiment two - online experiment. Magnitudes (standardised estimates or probabilities), error bars (95% CI) and significance (*$p$ < .05; ***$p$ < .001) of the fixed effects of the mixed models, with the vertical black line indicating null effect: a) standard estimates of the effects condition and injustice level on A-B dwell time (Model 2.1); b) probabilities of the effects of injustice level, A-B dwell time and condition on the likelihood to punish (note that, since punishment and compensation were the only two choices available, they are perfectly collinear, hence there is no compensation plot) (Model 2.2); c) standard

estimates of the effects injustice level, A-B dwell time and condition on the amount spent to punish and compensate (Model 2.3).

### 2.6 Alternative hypothesis 3 testing

As for experiment 1, we ran an alternative mixed model to partially test hypothesis 3, i.e., that the condition influences the amount spent. Specifically, we considered the predictive effects of injustice, difference in dwell time, condition, and response, on the amount spent. Results show that the higher the injustice the more participants would spend, and that participants spent more to punish (vs compensate), and spent more in the coin (vs digit) condition. No effect of difference in dwell time was found; however, as for experiment 1, difference in dwell time does not predict the amount spent per se, but the amount spent to punish (or compensate). Figure 3 shows the estimates, including error bars and p-values.


**Figure 3.** Experiment two – online experiment. Standardised estimates of the effects injustice level, Response (punish vs compensate), condition and A-B dwell time on the amount spent;

error bars (95% CI) and significance (*$p < .05$; ***$p < .001$) are also reported, with the vertical black line indicating null effect.

### 3. Exploratory Findings in Experiments One and Two: evidence of offender bias

Our variables of interest, i.e., attention, likelihood of reaction, and amount spent to react, were further analysed by comparing them on the type of reaction (punishment - compensation) and the target of the attention (offender - victim); this was done to obtain a more detailed picture of participants' preferences and understand whether one reaction (e.g., punishment) was preferred over the other (e.g., compensation), and whether one target (e.g., offender) was more attractive than the other (e.g., victim), irrespective of the mode of presentation (coins or digits). To do so, we performed a one sample t-test on the difference between the proportion of punishment and compensation (test value = 0 for the lab-based experiment; test value = 0.5 for the online experiment, which, as opposed to the lab-based experiment, only had two options). Both t-tests showed that, despite a larger difference in the coins condition, punishment was preferred to compensation in both conditions. A third one sample t-test (test value = 0) was performed on the difference between the dwell time on A and dwell time on B; similarly, it showed that, despite the difference being larger in the coins condition, participants preferred to look at the offender A rather than the victim B, in both conditions (see Table 3 and Figure 4 below for a summary of the statistics).

*Table 3. Exploratory findings for experiments one and two. a) One-sample t-tests for the lab-based experiment. A_B_Dwell refers to the difference between the time spent looking at A and B; prop_puncomp refers to the proportion of punishment minus compensation choices. b) One-sample t-tests for the online experiment. A_B_Dwell refers to the difference between the time spent looking at A and B; prop_pun refers to the proportion of punishment choices.*

| a) Experiment one (lab-based eye-tracking) | | | | | | 95% CI for mean difference | |
|---|---|---|---|---|---|---|---|
| **test value = 0** | **t(35)** | **p** | **Cohen's d** | **M** | **SD** | **Lower** | **Upper** |
| A_B_Dwell_coin | 4.29 | < .001 | 0.71 | 0.14 | 0.20 | 0.08 | 0.21 |
| A_B_Dwell_digit | 3.14 | 0.003 | 0.52 | 0.05 | 0.1 | 0.02 | 0.09 |
| Prop_puncomp_coin | 2.53 | 0.016 | 0.42 | 0.13 | 0.31 | 0.03 | 0.23 |
| Prop_puncomp_digit | 1.95 | 0.064 | 0.32 | 0.11 | 0.33 | -0.01 | 0.22 |

| b) Experiment 2 (online eye-tracking) | | | | | | 95% CI for mean difference | |
|---|---|---|---|---|---|---|---|
| **test value = 0** | **t(200)** | **p** | **Cohen's d** | **M** | **SD** | **Lower** | **Upper** |
| A_B_Dwell_coin | 6.49 | < .001 | 0.46 | 2.59 | 5.65 | 1.80 | 3.37 |
| A_B_Dwell_digit | 3.89 | < .001 | 0.27 | 1.57 | 5.71 | 0.77 | 2.36 |
| **test value = 0.5** | | | | | | | |
| Prop_pun_coin | 9.76 | < .001 | 0.69 | 0.67 | 0.24 | 0.13 | 0.20 |
| Prop_pun_digit | 8.53 | < .001 | 0.60 | 0.65 | 0.25 | 0.11 | 0.18 |

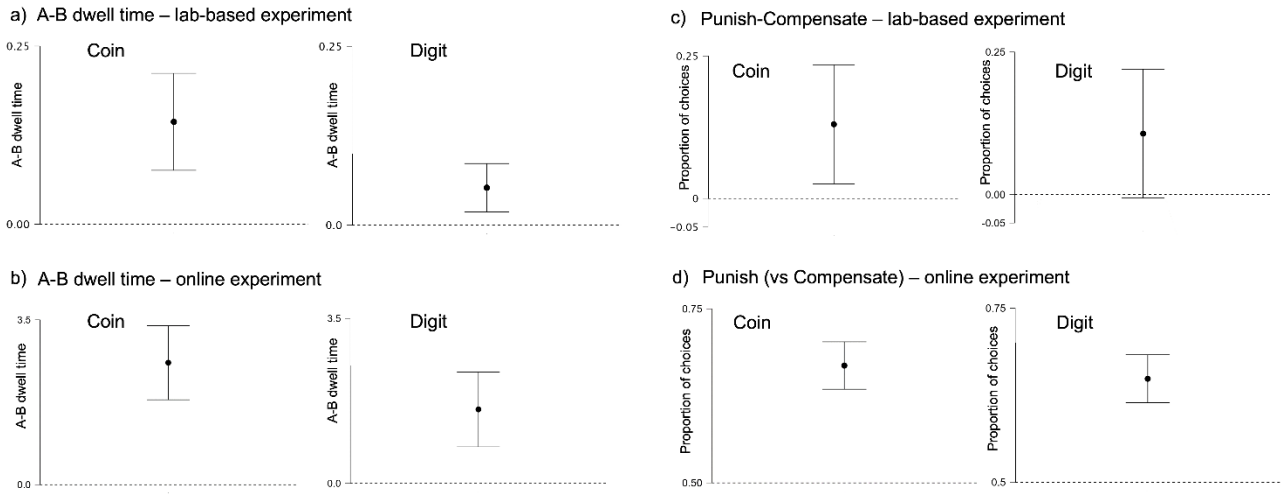


**Figure 4.** Exploratory findings, One-sample t-tests (95% CI) showing an offender bias and a punishment preference for each condition (coins and digits) on the: a) and b) A minus B dwell time, for experiment one (lab-based) and experiment two (online), respectively; c) proportion

of punishment minus compensation choices for experiment one, and d) proportion of punishment (vs compensation) choices for experiment two, where there were only two choices, and therefore the test-value (indifference point) is 0.5.

## 4. Experiment Three

### 4.1 Descriptive statistics.

*Table 4. Experiment three. Descriptive statistics for the variables in the TD experiment. Reveal_payoff_A refers to the proportion of offender's payoff being revealed; punish refers to the proportion of punishment choices; Amount refers to the average amount spent to punish or compensate; cognitive and affective empathy refer to the scores in the QCAE.*

| | Valid | Missing | M | SD | Min | Max |
|---|---|---|---|---|---|---|
| Reveal_payoff_A | 285 | 0 | 0.63 | 0.23 | 0.12 | 1 |
| Punish | 285 | 0 | 0.64 | 0.24 | 0 | 1 |
| Amount_punish | 283 | 2 | 46.79 | 22.89 | 0 | 100 |
| Amount_compensate | 239 | 46 | 36.95 | 21.72 | 0 | 100 |
| Cognitive_empathy | 285 | 0 | 2.91 | 0.42 | 1.44 | 3.94 |
| Affective_empathy | 285 | 0 | 2.71 | 0.51 | 1.33 | 4 |

### 4.2 Manipulation check.

As expected, the level of injustice positively predicted the amount spent to punish ($\beta = 0.4$; s.e. = 0.02, $t(1232) = 23.09$, $p < .001$) and compensate ($\beta = 0.32$; s.e. = 0.02, $t(666) = 13.23$, $p < .001$) (Figure 4b in the main text), indicating that the more severe the injustice, the more people spent to react to it, hence passing the sanity check.

## 5. Experiment Four

### 5.1 Descriptive statistics.

*Table 5. Experiment four. Descriptive statistics for the variables in the BU experiment. Punish_A and Punish_B refer to the proportion of punishment choices when A's and B's*

*payoffs were revealed; Amount refers to the average amount spent to punish or compensate; cognitive and affective empathy refer to the scores in the QCAE.*

| | Valid | Missing | M | SD | Min | Max |
|---|---|---|---|---|---|---|
| Punish_A | 284 | 0 | 0.74 | 0.29 | 0 | 1 |
| Punish_B | 284 | 0 | 0.48 | 0.32 | 0 | 1 |
| Amount_Punish_A | 267 | 17 | 45.90 | 24.73 | 0 | 100 |
| Amount_Compensate_A | 167 | 117 | 38.83 | 26.46 | 0 | 100 |
| Amount_Punish_B | 245 | 39 | 40.17 | 26.55 | 0 | 100 |
| Amount_Compensate_B | 240 | 44 | 37.53 | 23.14 | 0 | 100 |
| Cognitive_empathy | 284 | 0 | 2.87 | 0.43 | 1.59 | 4 |
| Affective_empathy | 284 | 0 | 2.72 | 0.53 | 1.25 | 3.83 |

### 5.2 Manipulation check.

As for experiment three (TD), the amount spent to punish ($\beta = 0.37$; s.e. = 0.02, $t(1137) = 20.73$, $p < .001$) and compensate ($\beta = 0.28$; s.e. = 0.02, $t(703) = 12.86$, $p < .001$) was positively predicted by the level of injustice, hence passing the sanity check (Figure 6b in the main text).

## 6. Experiment Five

### 6.1 Descriptive statistics.

*Table 6. Experiment five. Descriptive statistics for the variables in the BU task +Baseline experiment. Punish_Base refer to the proportion of punishment choices in the baseline condition (across coins and digits). Punish_A and Punish_B refer to the proportion of punishment choices when A's and B's payoffs were revealed; Amount refers to the average amount spent to punish or compensate; cognitive and affective empathy refer to the scores in the QCAE.*

| | Valid | Missing | M | SD | Min | Max |
|---|---|---|---|---|---|---|
| Punish_Base | 292 | 0 | 0.65 | 0.25 | 0 | 1 |
| Amount_Punish_Base | 284 | 8 | 46.38 | 25.58 | 0 | 100 |
| Amount_Compensate_Base | 262 | 30 | 36.71 | 24.73 | 0 | 100 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Punish_A | 292 | 0 | 0.75 | 0.29 | 0 | 1 |
| Amount_Punish_A | 276 | 16 | 43.75 | 24.80 | 0 | 100 |
| Amount_Compensate_A | 161 | 131 | 39.18 | 24.67 | 0 | 100 |
| Punish_B | 292 | 0 | 0.53 | 0.35 | 0 | 1 |
| Amount_Punish_B | 248 | 44 | 39.68 | 25.99 | 0 | 100 |
| Amount_Compensate_B | 225 | 67 | 36.77 | 22.72 | 0 | 100 |
| Cognitive_empathy | 292 | 0 | 55.60 | 8.29 | 26 | 76 |
| Affective_empathy | 292 | 0 | 32.49 | 6.01 | 14 | 47 |

### 6.2 Manipulation check.

This effect has been calculated considering both baseline and BU data. As for the previous experiments, the level of injustice positively predicted both the amount spent to punish ($\beta$ = 0.30; s.e. = 0.01, $t(4282) = 30$, $p < .001$) and the amount spent to compensate ($\beta = 0.31$; s.e. = 0.01, $t(2275) = 22.51$, $p < .001$), hence passing the sanity check.

## 7. Meta-analyses

Since the results were at times slightly inconsistent, we ran three mini meta-analyses synthesising the results of the eye-tracking experiments (1 and 2) and the choice and the severity effects across all experiments. We report the results below and also in the OSF, in the folder Mini Meta-analyses

https://osf.io/5egx8/?view_only=b8a20e04eeb345e486b19a58cd9ee224

### 7.1 Meta-analysis on the effect of attention on gaze

We ran, using JASP, a mini meta-analysis on the two eye-tracking experiments focusing on the effect of attention manipulation on gaze, considering the difference between the dwell time on the offender compared to the victim in the coin and the digit condition. A Restricted Maximum Likelihood model (RELM) was fitted to the data, and we interpreted the results

with caution, given that the test of residual heterogeneity was significant and very high (Q=5.24, p = 0.022; $I^2$ (%) = 80.9) and the number of studies very limited. We report the forest plot below: the overall effect was not significant (estimate = 0.33, se = 0.21, p = .123).

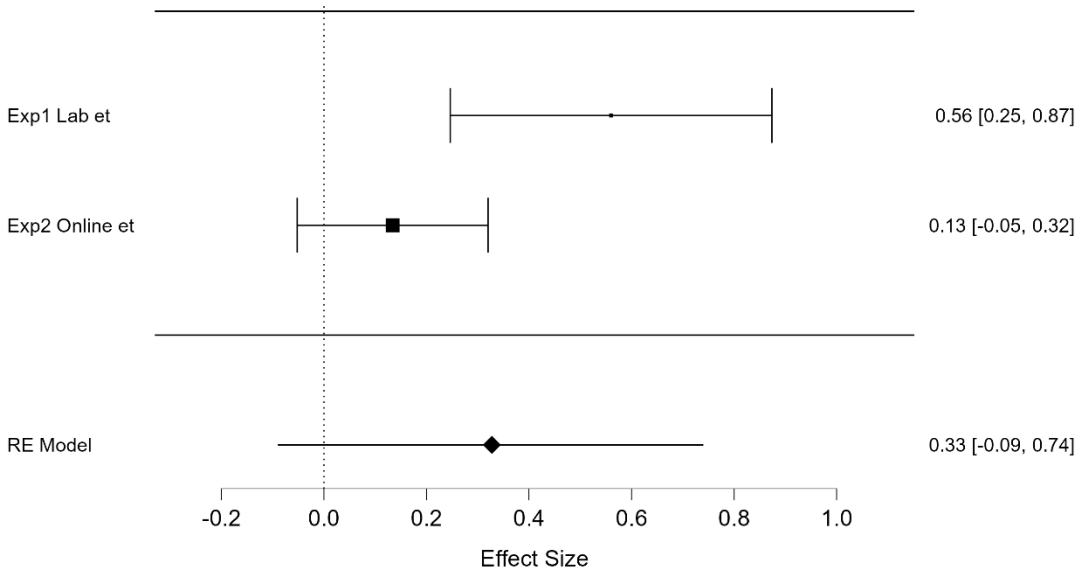


**Figure 4.** Effect of attention manipulation on gaze for the eye-tracking experiments. Standardised estimates of the effect sizes for each experiment; error bars (95% CI) are also reported. The diamond shape represents the overall effect size for the random effects model (weighted average of all studies).

### 7.2 Meta-analysis on the effect of choice

For the choice effect, we ran a RELM where we considered the effect size and its standard error (Cohen's d and se) of the t-tests comparing the choice to punish (vs compensate) when the offender is visually relevant vs when it is less relevant, across the 5 studies. This means that we considered the choice to punish (vs compensate) in the coin condition vs digit condition for the eye tracking experiments (1 and 2) and the choice to punish (vs compensate) when the offender (A) was revealed vs when the victim (B) was revealed in the information

frame experiments (3-5). We report the forest plot below: the overall effect was significant (estimate = 0.41, se = 0.10, $p < .001$), but we advise caution because the heterogeneity was very high ($Q = 54.78$, $p < .001$; $I^2$ (%) = 90.66).

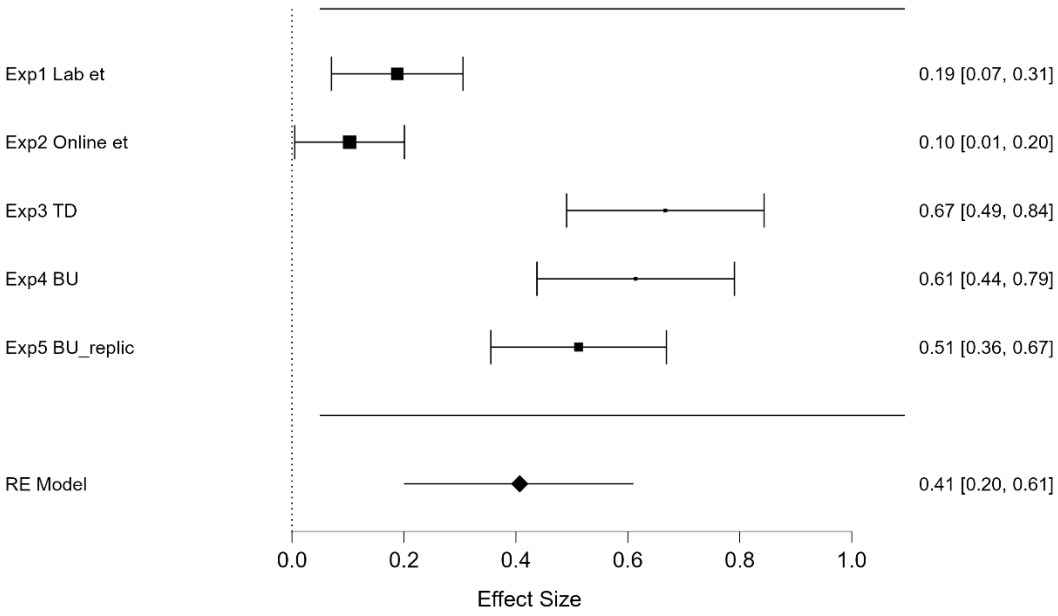


**Figure 5.** Effect of choice for all experiments. Standardised estimates of the effect sizes for each experiment; error bars (95% CI) are also reported. The diamond shape represents the overall effect size for the random effects model (weighted average of all studies).

### 7.3 Meta-analysis on the effect of severity (amount spent to punish)

For the severity effect, we ran a RELM where we considered the effect size and its standard error (Cohen's d and se) of the t-tests comparing the amount spent to punish when the offender is visually relevant vs when it is less relevant, across the 5 studies. This means that we considered the amount spent to punish in the coin condition vs digit condition for the eye tracking experiments (1 and 2) and the amount spent to punish when the offender (A) was revealed vs when the victim (B) was revealed in the information frame experiments (3-5). We report the forest plot below: the overall effect was significant (estimate = 0.29, se = 0.05, p <

.001). However, caution is advised due to the high heterogeneity (Q = 19.02, p < .001; $I^2$ (%) = 80.18).

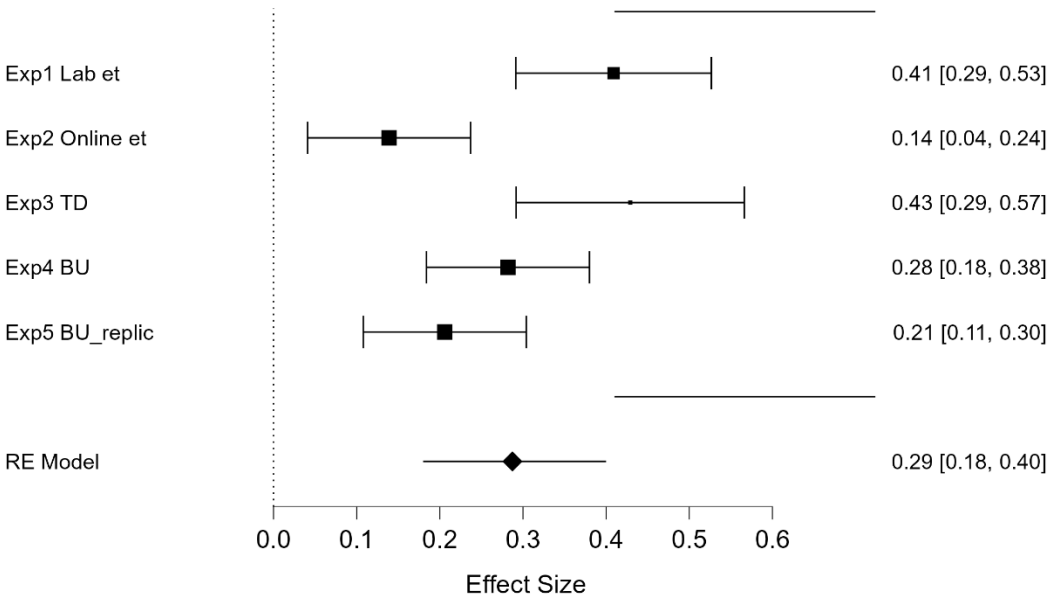


**Figure 6.** Effect of severity (amount spent to punish) for all experiments. Standardised estimates of the effect sizes for each experiment; error bars (95% CI) are also reported. The diamond shape represents the overall effect size for the random effects model (weighted average of all studies).