



Testing the Intuitive Retributivism Dual Process Model

Paul Rehren¹ and Valerij Zisman²

¹Ethics Institute, Department of Philosophy and Religious Studies, Utrecht University, The Netherlands

²Department of Philosophy, Faculty of History, Philosophy and Theology, Bielefeld University, Germany

Abstract: Research on the motives individuals have to punish criminal offenders suggests that punitive reactions are primarily driven by retributive, not utilitarian, motives. To explain this, several authors have suggested a dual process model (DPM) of punitive reactions. According to this model, punitive reactions are the product of two distinct types of processing (type I and type II), which differentially support retributive vs. utilitarian punishment motives. In response to cases of criminal wrongdoing, type I swiftly outputs a retributive reaction. In contrast, for utilitarian motives to play a role, this reaction has to be overridden by type II processing, which only happens rarely. In this article, we argue that despite its popularity, there is little concrete evidence for the DPM. We then report the results of a preregistered study investigating the effect of increased processing effort on retributive vs. utilitarian punitive reactions. We argue that the results fail to support the DPM.

Keywords: retributivism, utilitarianism, punitive reactions, dual process model, processing depth

When confronted with a case of criminal wrongdoing, most people's reaction will be that the culprit needs to be punished in some way (Henrich et al., 2006; Hoffman & Goldsmith, 2004). Despite its ubiquity, this punitive reaction may be supported by a number of different and sometimes conflicting motives. Inspired by a much longer standing debate in philosophy, psychological research into punishment motives has focused on two types of motives: retributivism and utilitarianism.¹ Retributive motives are backward-looking, meaning that they do not reference the future consequences of punishment. Instead, to punish out of retributive considerations is to think that offenders deserve punishment and that punishment is, therefore, an intrinsically appropriate response to wrongdoing (Duff & Hoskins, 2019; Kant, 1785/1998)

In contrast, utilitarian motives are forward-looking, focusing not on the past wrongdoing but the beneficial consequences of the punishment (Bentham, 1830/1998; Wood, 2010). While there are numerous proposals for what precisely the beneficial consequences of punishment are, psychologists have focused on two: deterrence – punishment deters the offender or other would-be criminals from committing similar offenses in the future; incapacitation – while the offender is undergoing punishment (e.g., incarceration), they will not be able to commit further crimes.

Previous literature paints an intriguing picture of why people punish. When asked directly, people by and large

report that both retributive and utilitarian considerations matter for punishment (for a review, Cullen et al., 2000). However, when punitive reactions are assessed by behavioral measures instead of self-report, retributive motives tend to dominate utilitarian motives (for a review, Editorial, this issue). This result has been dubbed the *intuitive retributivism hypothesis*.

The intuitive retributivism hypothesis has led several authors to propose a dual process model of punitive reactions (e.g., Aharoni & Fridlund, 2012; Darley, 2009; Keller et al., 2010). Dual process models have been proposed for a variety of domains of cognition. The central assumption of such models is that “cognitive tasks evoke two forms of processing that contribute to observed behavior” (Evans & Stanovich, 2013, p. 225), with the two forms – type I processing and type II processing – being qualitatively distinct. Typically, the distinction is thought roughly to line up with the more familiar distinction between intuition and deliberation. Different researchers differ considerably in how they spell out the distinctiveness of type I and type II processing (Evans, 2008, p. 200; Evans & Stanovich, 2013). However, common attributes of type I processing are that it is fast, parallel, automatic, and does not require working memory; in contrast, type II processing is often argued to be slow, serial, controlled, and to require working memory.

The most detailed presentation of the dual process model of punitive reactions is due to Carlsmith and Darley (2008).

¹ Some authors have investigated other motives, such as communication (e.g., Funk et al., 2014; Nadelhoffer et al., 2013) and restorative justice (for a review, van Doorn & Brouwers, 2017).

Carlsmith and Darley suggest that type I and type II processing differentially support different punishment motives. More specifically, they hypothesize that retributive reactions (punitive reactions responsive to retributive factors) are primarily the output of type I processing. In contrast, utilitarian reactions (punitive reactions responsive to utilitarian factors) require type II processing. When people are confronted with a case of criminal wrongdoing, type I processing engages and swiftly outputs a retributive reaction. This reaction can sometimes be overridden by type II processing, allowing for utilitarian motives to play a role. However, according to Carlsmith and Darley, this only happens rarely. Thus, in normal circumstances, people's punitive reactions are largely determined by an initial intuition that skews heavily retributive.

Problem

Carlsmith and Darley's dual process model (DPM) provides a neat explanation for the intuitive retributivism hypothesis: If punitive reactions tend to be the output of type I processing, which is primarily responsive to retributive factors and is only infrequently overwritten by more utilitarian type II processing, then people's punitive reactions should usually be retributive. This is indeed what the data suggests.

However, their model is not the *only* explanation for the data. For one thing, a single-process model like the rule-based account proposed by Kruglanski and Gigerenzer (2011) might also be able to capture the results. Even if we accept that a dual process framework is helpful for understanding punitive reactions, however, alternative explanations remain live. For instance, people may have both retributive and utilitarian intuitions, but the former tend to prevail. Alternatively, most initial type I punitive reactions may skew utilitarian but are then routinely overwritten by retributive type II processing.

The intuitive retributivism hypothesis *by itself* does not rule out any of these alternatives in favor of the DPM. Again, what the intuitive retributivism hypothesis suggests is that people's punitive reactions are primarily responsive to retributive, but not to utilitarian factors. This by itself, however, does not imply anything about the details of the psychological mechanism that underlies this pattern. In order to establish the DPM, then, additional work is required.

Some authors appear to be confident that this work has already been done. Darley (2009), for instance, describes the DPM as a "relatively clear picture of the naive psychology of punishment" (p. 2). Similarly, both Carlsmith and Darley (2008) and Robinson and Darley (2007) draw a number of policy implications from the DPM, suggesting that they, too, believe the model to be reasonably securely established.

At first glance, a number of experimental results do seem to directly support the DPM. First, Need for Cognition, an individual difference measure of cognitive style sometimes used to assess the tendency of individuals to engage in type II processing (Petty et al., 2009) has been found to be negatively associated with punitiveness (Sargent, 2004). Second, punitive reactions become more severe with cognitive load (Gollwitzer et al., 2016; Oswald & Stucki, 2009; van Knippenberg et al., 1999). Since type II processing requires cognitive resources (in particular, working memory capacity) to a much greater extent than type I processing, burdening those resources by inducing cognitive load is commonly used in experiments to inhibit type II processing (Evans & Stanovich, 2013, p. 232). Third, punitive reactions become less severe when participants are induced to think more carefully about their decision (Gollwitzer et al., 2016; Oswald & Stucki, 2009) – a manipulation that is thought to increase type II processing effort (Evans & Stanovich, 2013, p. 232).

However, all of these studies share one crucial limitation: In all of them, punitive reactions were only investigated generically – that is, without controlling for or looking at the underlying punishment motives. So while it is possible that the connection between type I processing and more severe punitive reactions reflects an association of *retributive* motives with type I processing, there is no way to tell. The pattern could instead be due to an association of *utilitarian* motives with type I processing or to an association of both retributive and utilitarian motives with type I processing. The same point holds for the connection between type II processing and less severe punitive reactions. In other words: From the fact that punitive reactions vary with processing type, little can be inferred about the psychology of the underlying punishment motives. For that, we also need to know which motives are associated with which processing type.

The second line of argument in support of the DPM points to research on dual-process models of moral cognition (Carlsmith & Darley, 2008; Darley, 2009). Two influential dual process models of moral cognition, Haidt's SIM (2001) and Greene's dual process model (2014), suggest that moral judgments tend to be intuitive in nature (the output of type I processing) and are only seldom overwritten by moral reasoning (type II processing). Carlsmith and Darley (2008, pp. 211–217) cite both models in support of their DPM, in effect suggesting that their model can be seen as a straightforward extension.

However, this strikes us as unpersuasive. Criticisms of dual process accounts of moral judgment aside (e.g., Kahane, 2012; Pizarro & Bloom, 2003), the argument assumes that punitive reactions can straightforwardly be treated as moral judgments. Yet it is not clear that this is the case. Many legal scholars argue that while there is

considerable overlap between the law and morality, there are nevertheless significant conceptual differences between the two domains (for discussion, Peczenik, 2005, chapter 4). More importantly, some research suggests psychological differences between judgments of punishment and paradigmatic moral judgments, like judgments about right and wrong or judgments about permissibility (Barbosa & Jiménez-Leal, 2017; Cushman, 2008; cf. Malle 2021, pp. 3.12–3.13). In contrast, while some researchers sometimes seem to lump the two types of judgment together (e.g., Greene, 2014, pp. 705–706), we know of no evidence that would justify this. In light of these points, we believe that the extent to which dual process models of moral cognition lend support to the DPM is questionable.

We think that the strongest evidence for the DPM comes from Aharoni and Fridlund (2012). Aharoni and Fridlund (Experiment 2) show that punitive reactions are susceptible to dumbfounding. Participants read a description of a crime for which the efficacy of common utilitarian motives for punishment had been minimized and were asked to recommend a punishment. Participants were then challenged to justify their decision. If a participant cited common utilitarian reasons, they were reminded that these considerations did not apply to the crime at hand. A majority of participants continued to recommend punishment even while admitting that no utilitarian reasons applied and not being able to articulate other reasons for their decision. Aharoni and Fridlund take this to suggest that people's punitive reactions are "shaped more by heuristic than rational processes" (p. 17).

Though suggestive, we think that the extent to which this finding supports the DPM is limited. First, the study was exploratory and small ($n = 47$). Second, not all participants were dumbfounded, leaving open the possibility that a subset of punitive reactions was responsive to utilitarian factors. Third, a similar argument has recently come under severe fire. Haidt and colleagues (2000) used semi-structured interviews to investigate people's reactions to harmless taboo violations. They report that for some violations, a majority of participants continued to maintain that the violation was wrong, even though they were unable to provide reasons for this. Many authors have cited these results in support of a dual process model of moral judgment (e.g., Haidt, 2001; Prinz, 2006). However, this move has repeatedly been challenged on methodological (Royzman et al., 2015) and conceptual grounds (Hindriks, 2015; Stanley et al., 2019). Due to the similarity of the designs of Aharoni and Fridlund and Haidt and colleagues, if successful, these objections would likely also call into question the extent to which dumbfounding can be appealed to in support of the DPM.

To summarize, while Carlsmith and Darley's (2008) DPM provides an appealing explanation for the intuitive

retributivism hypothesis, this alone is not enough to establish the DPM as a psychological fact. We have reviewed arguments in support of the DPM but have either found them unconvincing or have argued that they only provide limited support for the DPM. We conclude that some previous work (Carlsmith & Darley, 2008; Darley, 2009; Robinson & Darley, 2007) has overstated the extent to which the DPM is securely established.

Aims and Hypotheses

The first aim of this paper is to provide a direct test of the DPM. Recall that the DPM is a claim about which processes drive retributive versus utilitarian punitive reactions. Therefore, in order to test it, we will need two ingredients.

The first is a way of investigating punitive reactions that measures the underlying punishment motives. We here use the information search task approach of Keller and colleagues (2010, Experiment 2). Keller and colleagues put participants in charge of assigning punishment to an offender guilty of a crime. To inform their decision, participants were asked to select five items of information about the crime or the offender from a list. The items were chosen such that each would be relevant either from the retributive or from one of two utilitarian points of view (deterrence, incapacitation). For example, because retributivism is focused on past wrongdoing, features like the magnitude of harm that an offender has caused and the seriousness of the offense should mainly play a role in punitive reactions if the underlying motives are retributive. Conversely, features like the general frequency of a crime (deterrence) and the risk of offender recidivism (incapacitation) should primarily factor into utilitarian punitive reactions because utilitarianism focuses on the consequences of punishment.

The measure of interest was the order in which participants requested retributive items, deterrence items, and incapacitation items. To capture this order, Keller and colleagues calculated a rank-preference score (RPS) for each participant and punishment motive. Each item selection trial was weighted. The first trial received a weight of 5; the second trial received a weight of 4; and so on. For a given punishment motive, its rank-preference score was then calculated as the sum of the trial weights in which an item related to that motive was selected. For example, if a participant chose retributive items on the first, third, and fourth trials, their retributivism RPS would be $5 + 3 + 2 = 10$.

The second ingredient is the way of manipulating the type of processing participants rely on (for an overview, Horstmann et al., 2010). Here, we chose to increase type II processing effort by inducing some participants to think carefully about their tasks. This manipulation is commonly used in research on dual-process models (e.g., Evans et al., 2010; Oswald & Stucki, 2009; Shenhav et al., 2012).

Putting the two ingredients together, in our study, all participants completed the information search task from Keller and colleagues (2010, Experiment 2) while being randomly assigned to one of two conditions. In the treatment condition (Think Carefully), participants were induced to think carefully about which pieces of information they would request. To this end, we instructed them to make each request only after thorough deliberation and to take their time. In the control condition (Control), participants were not given any special instruction.

According to the DPM, punitive reactions are the product of two distinct types of processing (type I and type II). Type I processing tends to output retributive, punitive reactions. In contrast, for utilitarian motives to play a role, this reaction has to be overridden by type II processing. Thus, to the extent that our manipulation is successful in increasing type II processing effort, the DPM predicts that more participants in the Think Carefully condition than in the Control condition will override their initial retributive intuitions in favor of utilitarian punitive reactions. Therefore, the importance of retributive items should decrease relative to the Control condition, while the importance of deterrence and incapacitation items should increase (in our preregistration, these hypotheses are labeled H2a–H2c):

Hypothesis 1a (H1a): Participants in the Control condition will have higher retributivism rank-preference scores than participants in the Think Carefully condition.

Hypothesis 1b (H1b): Participants in the Control condition will have lower deterrence rank-preference scores than participants in the Think Carefully condition.

Hypothesis 1c (H1c): Participants in the Control condition will have lower incapacitation rank-preference scores than participants in the Think Carefully condition.

The second aim of this paper is to provide a (further) test of the intuitive retributivism hypothesis. Because the DPM was explicitly introduced in order to explain the intuitive retributivism hypothesis, it predicts that hypothesis. Recall that according to the intuitive retributivism hypothesis, under normal conditions, punitive reactions are driven more strongly by retributive than by utilitarian punishment motives. In the context of the information search task of Keller and colleagues (2010, Experiment 2) that we use here, the intuitive retributivism hypothesis thus predicts that participants in our Control condition (which does not include any additional instructions) will prioritize requesting

retributive items over deterrence items and incapacitation items. In other words (in our preregistration, these hypotheses are labeled H1a and H1b):

Hypothesis 2a (H2a): Participants in the Control condition will have higher retributivism rank-preference scores than deterrence rank-preference scores.

Hypothesis 2b (H2b): Participants in the Control condition will have higher retributivism rank-preference scores than incapacitation rank-preference scores

Methods

The study was preregistered (<https://dx.doi.org/10.23668/psycharchives.3479>). We deviated from our preregistration in our data collection and analysis. We explain and justify these deviations transparently in our methods and analysis sections. Our study materials and raw data (including codebook) are available as Electronic Supplementary Material (ESM) (Rehren & Zisman, 2021; see Open Data section at the end of this article).

Design

We asked participants to read a short text informing them that a crime had been committed, that the offender had been caught, and that their task would be to assign a punishment. In Keller and colleagues (2010, Experiment 2), all participants read about a residential burglary. In our study, the type of crime was randomly chosen from a set of five types of crime (blackmail, stolen property, arson, aggravated assault, murder). We included this variation on Keller et al. to improve generalizability.

Following this text, participants saw a list of short descriptions of pieces of information about the crime and the offender. The order of this list was randomized. Participants were asked to request items of information from this list in order to help them make their punishment decision. Participants could only choose one item at a time. In order to request an item, participants clicked on it, followed by clicking on a button labeled “Select item” below the list. Participants were instructed to request items in order of priority and were not made aware of how many items they would be able to request in total. This selection procedure was repeated six times.

The study was an experimental study using randomized control trials. Participants were randomly assigned to one of two conditions. In the Think Carefully condition, participants were asked to think carefully about each of

Table 1. Items of information used in our study

Items relating to retributivism

- (r1) *Seriousness of the offense*: How serious is this particular offense?
- (r2) *Magnitude of harm*: How severe is the financial, physical, or psychological harm that the offender has caused?
- (r3) *Offender intent*: Did the offender act with intention?
- (r4) *Extenuating circumstances*: Are there aspects of this particular offense that make the offender less than fully responsible?

Items relating to deterrence

- (d1) *Publicity of offense and trial*: Will this particular offense and its trial attract a lot of public attention?
- (d2) *Detection rate*: How frequently are offenses like this detected and brought to trial?
- (d3) *Frequency of the offense*: How frequently do offenses like this occur in society?
- (d4) *Frequency trend*: Is the frequency of offenses like this in society increasing or decreasing?

Items relating to incapacitation

- (i1) *Risk of offender recidivism*: How likely is it that the offender will commit further offenses?
- (i2) *Dangerousness of the offender*: How dangerous is the offender?
- (i3) *Self-control*: Does the offender normally have good self-control or does the offender frequently act on impulse?
- (i4) *Prior record*: Does the offender have a prior criminal record?

their requests. To this end, participants were instructed to only request an item after thorough deliberation and to take their time for each request. In contrast, participants in the Control condition did not receive any additional instructions.

Once participants had selected five items, they were shown the pieces of information about the crime and the offender that they had requested. Participants then indicated their punishment decision. To end the survey, participants provided standard demographic information, were thanked for their participation and exited the survey.

Materials

Item Descriptions

Our item descriptions differ from Keller and colleagues (2010, Experiment 2) in three ways. First, we included a number of additional items. We hoped that this would help us provide a more general test of the DPM. All of the additional items describe features of the crime or of the offender that have been used in previous research to probe retributive vs. utilitarian punishment motives. Second, we made minor changes to the wording of some of the original items to make them more precise. Third, we did not include the three filler items that were used by Keller and colleagues. Table 1 shows the items. Items (r1)–(r4) relate to retributivism. The remaining items relate to utilitarianism, with items (d1)–(d4) relating to deterrence and items (i1)–(i4) relating to incapacitation.

Once a participant had requested five items and before making their punishment decision, they received the items of information they had requested. For example, if the participant had requested offender intent, they were told that the offender had been planning the offense for several days.

If the participant had requested information about the risk of offender recidivism, they were informed that the offender had publicly stated that they would repeat their offense if given a chance. The exact wording of all pieces of information is provided in ESM 1.

Punishment

We used three items to measure punitive reactions. Because participants' punitive reactions are not the focus of our analyses, we do not describe this measure in detail here (see ESM 1).

Manipulation Check

For each item request, we recorded the time from first viewing the list of items to clicking "Select item." We expected longer item request times in the Think Carefully condition than in the Control condition (Horstmann et al., 2010).

Demographic Questions

Participants were asked to report their age, gender, ethnicity, religiosity, and political attitudes. Moreover, participants were asked whether they had ever taken an ethics course and a law course.

Exclusion of Participants

There were no participants with missing data. In order to reduce the likelihood that participants who did not read the study instructions conscientiously would end up in our analyses, we used an instructional attention check (Oppenheimer et al., 2009). As part of the study instructions, participants were told to select a specific item

(*Gender of the offender*: “What is the offender’s gender?”) on the first trial. We excluded the data of all 218 participants who failed to select this item from all analyses.

Participants

A simulation-based power analysis indicated that we would need a sample size of $n > 485$ to achieve the power of at least 90% (Chambers et al., 2019) when testing our hypotheses. We describe this power analysis in detail in ESM 5. The R code used to run the power analysis is available in ESM 6. To allow for exclusions due to attention check failure, we recruited a total sample of $n = 560$ (40.2% female; M (SD) = 55.6 (33.3) years; 5.0% Asian, 1.8% Black/African, 0.9% Caribbean, 90.0% Caucasian/White, 2.0% Other, 0.4% Prefer not to say). In our preregistration, we stated that we would recruit a total sample of $n = 559$. However, due to an error with our survey platform, data from one more participant were collected. One hundred ninety-eight participants started the survey but did not complete all study materials and so were not included in the total sample. A total of 218 participants failed the attention check, and were excluded.² The final sample thus consisted of $n = 342$ participants (47.1% female; M (SD) = 55.3 (13.8) years; 2.9% Asian, 1.5% Black/African, 0.9% Caribbean, 93.3% Caucasian/White, 0.9% Other, 0.6% Prefer not to say).

Procedure

Data were collected online through the ZPID’s PsychLab online (<https://leibniz-psychology.org/en/services/data-collection/psychlab-online/>) who purchased the sample from the respondi (<https://www.respondi.com>). Participants were UK residents who had registered through the respondi. In our preregistration, we stated that individuals would be considered eligible for participation if their first language was English and they had at least a 95% approval rate on the previous submission. However, since it was not possible to select participants based on these criteria through the respondi, we instead used UK residents (whose first language would typically be English) and dropped the first selection criterion.

Eligible participants received an invitation email containing a link to the study. Upon accepting to participate, participants were redirected to the study, which was hosted on LimeSurvey (<https://www.limesurvey.org>). After giving informed consent, participants read the study instructions and completed the study materials. In the end, participants

were thanked for their participation, exited the survey, and were compensated for their participation.

The respondi provides incentives for participation in the form of tokens or bonus points which, after a certain amount has been accumulated, participants can either have paid out to them or donate. The study was approved by the Duke University Campus Institutional Review Board in accordance with the declaration of Helsinki. The data are available in ESM 2.

Analysis

All analyses were carried out in R (R Core Team, 2020). The analysis code is provided in ESM 4.

Descriptive Statistics

Table 2 shows the number of participants who requested an item corresponding to a given punishment motive (retributivism, deterrence, incapacitation) on a given trial. In addition, we calculated means and standard deviations for each rank-preference score (RPS), as well as pairwise Spearman’s rank-order correlations between the different RPS (Table 3).

Preregistered Analysis

To check our manipulation, we entered condition (Control, Think Carefully) into a linear mixed-effects model predicting item selection time (Bates et al., 2015). The model was fit using ML. P -values were computed using Satterthwaite’s approximation of denominator degrees of freedom (Kuznetsova et al., 2017). We added a random intercept for the participant. In our preregistration, we had planned to add an additional random intercept for the type of crime (blackmail, stolen property, arson, aggravated assault, murder); however, this model did not converge. As expected, participants in the Think Carefully condition ($M = 31.6$ s, $SD = 32.6$ s) spend more time on item requests than participants in the Control condition ($M = 27.6$ s, $SD = 28.8$ s), $b = -3.93$, $SE = 1.68$, $df = 340.0$, $t = -2.34$, $p = .0201$.

To test H1a–H2b, we performed a planned contrast analysis (Schad et al., 2020). We first combined condition and motive into one factor with six levels. We then expressed the nulls corresponding to our five hypotheses in terms of contrasts of group means indexed by the levels of this factor. We extracted the contrast coefficients and combined

² The results of none of the analyses reported in this paper change significantly (i.e., statistically significant effects remain significant, statistically nonsignificant effects remain nonsignificant) when we include these participants (see ESM 6).

Table 2. Number of participants who requested an item related to a given punishment motive on a given trial, by experimental condition

Trial	Condition	Retributivism					Deterrence					Incapacitation				
		(r1)	(r2)	(r3)	(r4)	Total	(d1)	(d2)	(d3)	(d4)	Total	(i1)	(i2)	(i3)	(i4)	Total
1	Think Carefully	40	33	25	14	112	0	0	1	0	1	7	26	2	30	65
	Control	30	48	25	13	116	0	2	0	0	2	3	17	2	24	46
2	Think Carefully	26	34	36	15	111	0	0	1	1	2	10	27	3	25	65
	Control	41	26	23	17	107	1	0	0	1	2	5	21	7	22	55
3	Think Carefully	21	27	33	19	100	1	0	0	1	2	16	30	5	25	76
	Control	21	27	23	16	87	0	0	1	2	3	18	27	7	22	74
4	Think Carefully	25	25	24	20	94	0	0	3	6	9	19	24	15	17	75
	Control	14	22	27	17	80	1	1	2	2	6	18	29	8	23	78
5	Think Carefully	23	17	20	25	85	2	3	2	2	9	22	19	12	31	84
	Control	16	17	24	19	76	1	0	1	0	2	25	26	10	25	86

Table 3. RPS means, standard deviations and pairwise Spearman's rank-order correlations, by experimental condition; square brackets show 95% confidence intervals (CIs)

Motive	Condition	<i>M</i>	<i>SD</i>	<i>r</i>						
				1.	2.	3.	4.	5.	6.	
1. Retributivism	Think Carefully	8.86	3.14	-		-0.21 [- 0.33, -0.09]			-0.96 [- 0.98, -0.93]	
	Control	9.18	2.69	-			-0.14 [- 0.28, 0.00]			-0.96 [- 0.98, -0.93]
3. Deterrence	Think Carefully	0.26	0.91			-			-0.03 [- 0.17, 0.09]	
	Control	0.25	0.87							-0.11 [- 0.24, 0.04]
5. Incapacitation	Think Carefully	5.88	3.08							
	Control	5.57	2.70							

them into a contrast matrix. We describe this procedure in detail in ESM 5. Next, we entered the contrasts into a linear mixed-effects model predicting RPS. We added a random intercept for the participant. Again, the preregistered model, which included an additional random intercept for the type of crime, did not converge.

The model supports h2a and h2b. For participants in the Control condition, there was a significant difference between retributivism RPS ($M = 9.18$, $SD = 2.69$) and deterrence RPS ($M = 0.25$, $SD = 0.87$), $b = 8.93$, $SE = 0.27$, $df = 1,020.0$, $t = 33.15$, $p < .001$, and between retributivism RPS and incapacitation RPS ($M = 5.57$, $SD = 2.70$), $b = 3.60$, $SE = 0.27$, $df = 1,020.0$, $t = 13.38$, $p < .001$.

To compare these results with the results obtained by Keller and colleagues (2010, Experiment 2), we calculated Hedges' g_{av} (Lakens, 2013) for the mean differences between retributivism RPS and deterrence RPS, and retributivism RPS and incapacitation RPS for the two studies (Table 4). Table 4 shows that the g_{av} we obtained were larger than (but comparable to) the g_{av} obtained by Keller and colleagues.

In contrast, we did not find support for the other three hypotheses (H1a-H1c). Retributivism RPS did not

significantly differ between the Think Carefully condition ($M = 8.86$, $SD = 3.14$) and the Control condition ($M = 9.18$, $SD = 2.69$), $b = 0.32$, $SE = 0.27$, $df = 1,020.0$, $t = 1.20$, $p = .230$. The same was true for deterrence RPS ($M = 0.26$, $SD = 0.91$ vs. $M = 0.25$, $SD = 0.87$), $b = 0.01$, $SE = 0.27$, $df = 1,020.0$, $t = 0.03$, $p = .975$, and incapacitation RPS ($M = 5.88$, $SD = 3.08$ vs. $M = 5.57$, $SD = 2.70$), $b = 0.31$, $SE = 0.27$, $df = 1,020.0$, $t = 1.17$, $p = .242$. Figure 1 illustrates the results.

Additional Analysis

While participants in the Think Carefully condition did spend longer on their item requests than participants in the Control condition, this difference was quite small (3.93 s). This suggests the possibility that our manipulation may have been too subtle to reveal an effect of processing type on punishment motives. Luckily, our study allows for a second way of probing the relationship between processing type and the motives underlying our participants' punitive reactions. To the extent that participants who spend longer on their item requests tend to rely more on type II

Table 4. Hedges' g_{av} for the mean differences between retributivism rank-preference score (RPS) and deterrence RPS, and retributivism RPS and incapacitation RPS for Keller and colleagues (2010, Experiment 2) and our control condition

	Hedges' g_{av}	
	Keller and colleagues (2010, Experiment 2) <i>n</i> = 78	Current study <i>n</i> = 164
Retributivism vs. deterrence RPS	3.38	4.45
Retributivism vs. incapacitation RPS	1.26	1.33

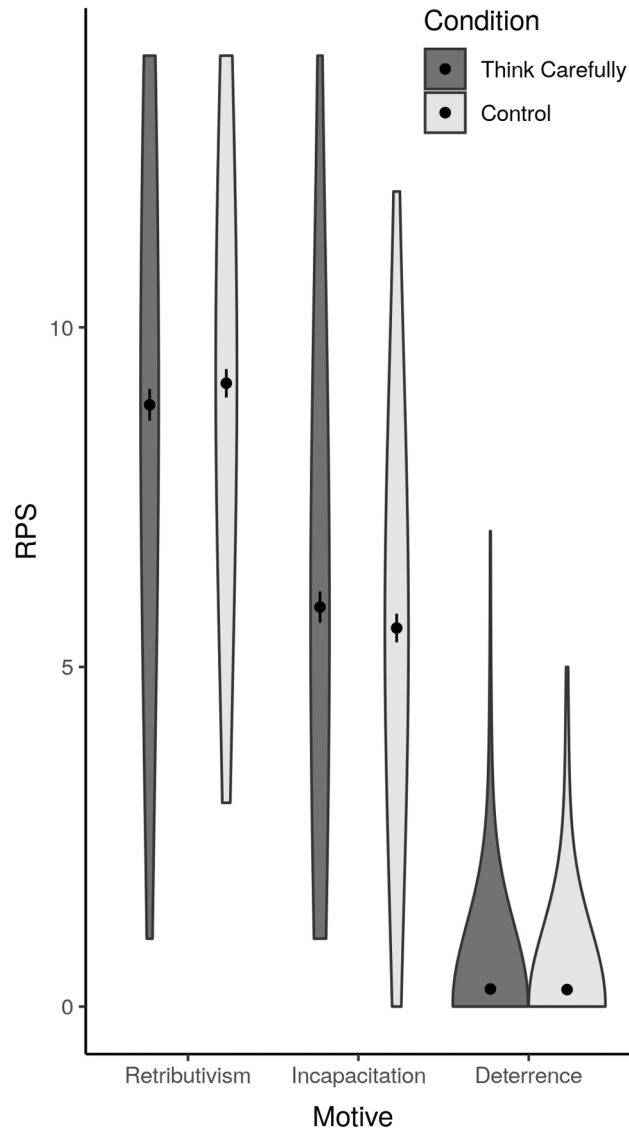


Figure 1. Violin plot of retributivism, deterrence, and incapacitation rank-preference score (RPS) by condition. Black dots show means, error bars show standard errors.

processing and less on type I processing (cf. Evans, 2008), the DPM predicts that item request time would be negatively associated with the importance of information related to retributivism, and positively associated with the importance of both information related to deterrence and information related to incapacitation. To investigate this,

we calculated Spearman's rank-order correlations between item request time on one hand and retributivism RPS, deterrence RPS, and incapacitation RPS on the other. All of these correlations were very small, and none of them reached statistical significance, $|r| < 0.03$, $p > .310$. Thus, again, we failed to find evidence for the DPM.

Discussion

According to the intuitive retributivism hypothesis, punitive reactions are driven more strongly by retributive than by utilitarian punishment motives. The first aim of our study was to test this hypothesis. To do this, we used the information search task approach from Keller and colleagues (2010, Experiment 2). We asked participants to request pieces of information about a crime that we put them in charge of assigning punishment for. The items were chosen such that each item would be relevant either from the retributive or from one of two utilitarian points of view (deterrence, incapacitation). If the intuitive retributivism hypothesis is correct, we would expect participants to prioritize requesting retributive items over deterrence items and incapacitation items.

Our Control condition conceptually replicates Keller and colleagues (2010, Experiment 2), supporting the intuitive retributivism hypothesis: participants indeed requested more retributive items of information than both deterrence items and incapacitation items (see also Carlsmith, 2006). Both effects were larger in our study than in Keller and colleagues. Moreover, while our Control condition followed the experimental design of Keller and colleagues quite closely, it deviates in three important respects, all of which support the robustness of their results, and therefore of the intuitive retributivism hypothesis. First, we used participants from a different set of countries (UK vs. Switzerland, Germany, or Austria). Our findings, therefore, add to a growing list of (Western) countries where people's punitive reactions seem to be more strongly driven by retributive motives than by utilitarian motives. Second, our study included additional items of information, suggesting that a preference for retributive information over utilitarian information holds even when people are given more non-retributive options to choose from (cf. Keller et al., 2010,

https://econtent.hogrefe.com/doi/pdf/10.1027/2151-2604/a000461 - Friday, April 01, 2022 8:01:52 AM - IP Address: 82.72.253.231

Experiment 3). Third, we varied the type of crime that participants read about between participants, including crimes of different levels of severity. This suggests that people prefer retributive over utilitarian information for a wider variety of crimes than just residential burglaries (the crime used by Keller et al., 2010).

The most prominent explanation for the intuitive retributivism hypothesis in the literature is the dual-process model of punitive reactions (DPM; Aharoni & Fridlund, 2012; Darley, 2009; Keller et al., 2010; Carlsmith and Darley 2008). The DPM seeks to explain the intuitive retributivism hypothesis by suggesting that punitive reactions are the product of two distinct types of processing (type I and type II). Type I processing primarily outputs retributive punitive reactions. In contrast, utilitarian punitive reactions require type II processing to come online, which happens only infrequently. Thus, the DPM predicts that for individuals who are making a punishment decision, increased type II processing effort would decrease the importance of retributivism-related information and increase the importance of utilitarianism-related information.

The second aim of this study was to provide the first direct test of the DPM. Our study failed to find evidence for the DPM. Participants who were asked to think carefully about each item of information request, and to take their time, did not give higher priority to requesting deterrence items or incapacitation items over retributive items than participants who received no further instructions. Yet to the extent that being asked to think carefully and take one's time encourages type II processing, this is what the DPM predicts. Taken together with the lack of other strong arguments in favor of the DPM, we think that our results warrant a healthy amount of skepticism towards the DPM. Certainly, we believe that it is too early to describe the DPM as a "relatively clear picture of the naive psychology of punishment" (Darley, 2009, 2) or to draw policy implications from it (Carlsmith & Darley, 2008; Robinson & Darley, 2007).

Limitations

Our study is subject to several limitations. First, we used an online convenience sample of UK participants. This means that it is unclear whether our results generalize to people in other parts of the world (Henrich et al., 2010), or to the broader population of the UK.

Second, more participants failed our attention check and thus had to be excluded from analysis than we had estimated when calculating our sample size. Because of this, we did not reach the sample size required to achieve 90% power for our hypotheses tests. While a simulation-based sensitivity power analysis indicates that our design was sensitive enough to detect the smallest effect sizes of

interest that we had specified in our original power analysis (see ESM 6) with the power of at least 90%, this nonetheless reduces the level of confidence one should have in our findings.

Third, like much other research on what punishment motives people have, our study relies on the assumption that certain features of a crime are (more or less) uniquely relevant to retributivism or utilitarianism. This assumption has not gone unquestioned. In particular, Goodwin and Benforado (2015) have argued that many supposedly retributive features (e.g., magnitude of harm, offender intent) also matter from the perspective of deterrence. Goodwin and Benforado suggest that this constitutes a large confound for the interpretation of any research into the motives underlying punitive reactions that relies on this assumption. While we think that Goodwin and Benforado overstate their case somewhat (see ESM 7), we nonetheless agree that they raise an important conceptual worry. Hence, future research on the DPM should consider approaches that do not rely on this assumption (e.g., Goodwin & Benforado, 2015; McFatter, 1982).

Similarly, our study relies on the assumption that being related to different punishment motives is the only relevant systematic difference between our items. However, there may be other such differences. If so, then participants might prefer one type of information for reasons that are unrelated to what punishment motive is driving their punitive reactions. In particular, we think that it is plausible that some of our deterrence items are less straight-forwardly related to the punishment motive of deterrence than some of our retributivism items are related to retribution. To arrive at the conclusion that a given item of information is relevant from the point of view of deterrence, participants might need to go through more steps, or going through the required steps might take more effort, than for items related to retributivism. One intriguing possibility is that instead of being a confound, this (partially) explains the intuitive retributivism hypothesis: people prefer information related to retributivism because that information just feels more immediately relevant to the question of how much punishment an offender should receive. We think that future research should investigate this possibility. In the absence of such research, however, the point does potentially confound the interpretation of our results as supporting the intuitive retributivism hypothesis (at least regarding the comparison between retributivism and deterrence items). One way for future research to do better would be to use pre-tests in order to select items that do not differ systematically in how straight-forwardly they are related to their respective punishment motive.

Finally, while participants in the Think Carefully condition did spend more time on item requests than participants in the Control condition, this difference was small enough

to suggest that our manipulation might have been too subtle to reveal an effect of processing type on punishment motives. This reduces our confidence in our rejection of the DPM. We think that this reduction is somewhat moderated by our failure to find any correlations between the item request time and the type of information requested. This is because to the extent that participants who spend longer on their item requests tend to rely more on type II processing and less on type I processing, the DPM predicts that longer item request times will be associated with an increased preference for utilitarian information and a decreased preference for retributive information. Nevertheless, to fully overcome these doubts, future research on the DPM should consider making use of other, less subtle ways of manipulating processing type. Options include inducing cognitive load, working with time constraints, manipulating mood, or a combination of approaches.

Directions for Future Research

Our study suggests several directions for future research. First, previous studies have employed a variety of other approaches to probe people's punishment motives beyond the information search task we used here (see Editorial, this issue). Most of these studies also find support for the intuitive retributivism hypothesis. Therefore, it would be useful for future research to investigate the effect of manipulating processing type on the motives driving people's punitive reactions when measured in these other ways.

Second, while the intuitive retributivism hypothesis is often framed in terms of the relative importance of retributive vs. utilitarian punishment motives, some researchers have sought to broaden this scope to include other types of motives, for example, motives relating to restorative justice or to the communicative dimension of punishment. A number of studies suggest that retributive motives dominate in these contests, as well (e.g., Crockett et al., 2014; Gromet & Darley, 2009; Nadelhoffer et al., 2013; van Prooijen, 2010). If so, then this suggests a natural extension to the DPM: type II processing needs to come online and to overwrite retributive type I processing not only in order for utilitarian punishment motives to play a role in people's punitive reactions but also for communicative and reparative motives to play a role. Future research might test this broader version of the DPM.

Finally, if it is not the DPM that explains the intuitive retributivism hypothesis, then what does? Given the robustness of this hypothesis, it does seem reasonable to expect psychologists to come up with a plausible explanation for it. One idea could be to extend the DPM by adding a third type of reflective processing (Sauer, 2018; Stanovich, 2009). Some evidence links reflective processing to utilitarian moral judgment (e.g., Patil et al., 2021; Paxton

et al., 2012); hence, this theoretical perspective might help to explain when utilitarian motives play a role in punitive reactions, as well. Another promising direction is to dissolve the strong dichotomy between type I and type II processing. For example, recently, a number of theorists (e.g., Crockett, 2013; Cushman, 2013) have proposed rational learning approaches to dual-process models of moral judgment and decision-making, which highlight how episodes of moral learning (i.e., type II processing) can feedback into and shape type I processes. It could be useful to apply this approach to the DPM of punitive reactions.

References

- Aharoni, E., & Fridlund, A. J. (2012). Punishment without reason: Isolating retribution in lay punishment of criminal offenders. *Psychology, Public Policy, and Law*, 18(4), 599–625. <https://doi.org/10.1037/a0025821>
- Barbosa, S., & Jiménez-Leal, W. (2017). It's not right but it's permitted: Wording effects in moral judgement. *Judgment and Decision Making*, 12(3), 308–313.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bentham, J. (1998). The rationale of punishment. In T. Vormbaum (Ed.), *Strafrechtsdenker der Neuzeit* [A new era of criminal justice thinking] (pp. 387–394). Berliner Wissenschafts-Verlag. (Original work published 1830).
- Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology*, 42(4), 437–451. <https://doi.org/10.1016/j.jesp.2005.06.007>
- Carlsmith, K. M., & Darley, J. M. (2008). Psychological aspects of retributive justice. *Advances in Experimental Social Psychology*, 40, 193–236. [https://doi.org/10.1016/S0065-2601\(07\)00004-4](https://doi.org/10.1016/S0065-2601(07)00004-4)
- Chambers, C., Banks, G. C., Bishop, D., Bowman, S., Button, K., Crockett, M., Dienes, Z., Errington, T., Fischer, A., & Holcombe, A. O. (2019). *Registered reports*. <https://osf.io/8mpji/>
- Crockett, M. J. (2013). Models of morality. *Trends in Cognitive Sciences*, 17(8), 363–366. <https://doi.org/10.1016/j.tics.2013.06.005>
- Crockett, M. J., Özdemir, Y., & Fehr, E. (2014). The value of vengeance and the demand for deterrence. *Journal of Experimental Psychology: General*, 143(6), 2279–2286. <https://doi.org/10.1037/xge0000018>
- Cullen, F. T., Fisher, B. S., & Applegate, B. K. (2000). Public opinion about punishment and corrections. *Crime and Justice*, 27, 1–79. <https://doi.org/10.1086/652198>
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2), 353–380. <https://doi.org/10.1016/j.cognition.2008.03.006>
- Cushman, F. (2013). Action, outcome, and value: A dual-system framework for morality. *Personality and Social Psychology Review*, 17(3), 273–292. <https://doi.org/10.1177/1088868313495594>
- Darley, J. M. (2009). Morality in the law: The psychological foundations of citizens' desires to punish transgressions. *Annual Review of Law and Social Science*, 5(1), 1–23. <https://doi.org/10.1146/annurev.lawsocsci.4.110707.172335>
- Duff, A., & Hoskins, Z. (2019). Legal punishment. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/win2019/entries/legal-punishment/>

- Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255–278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Evans, J., Handley, S. J., Neilens, H., & Over, D. (2010). The influence of cognitive ability and instructional set on causal conditional inference. *The Quarterly Journal of Experimental Psychology*, 63(5), 892–909. <https://doi.org/10.1080/17470210903111821>
- Evans, J., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3), 223–241. <https://doi.org/10.1177/1745691612460685>
- Funk, F., McGeer, V., & Gollwitzer, M. (2014). Get the message: Punishment is satisfying if the transgressor responds to its communicative intent. *Personality and Social Psychology Bulletin*, 40(8), 986–997. <https://doi.org/10.1177/0146167214533130>
- Gollwitzer, M., Braun, J., Funk, F., & Süßenbach, P. (2016). People as intuitive retaliators: Spontaneous and deliberate reactions to observed retaliation. *Social Psychological and Personality Science*, 7(6), 521–529. <https://doi.org/10.1177/1948550616644300>
- Goodwin, G. P., & Benforado, A. (2015). Judging the goring ox: Retribution directed toward animals. *Cognitive Science*, 39(3), 619–646. <https://doi.org/10.1111/cogs.12175>
- Greene, J. D. (2014). Beyond point-and-shoot morality: Why cognitive (neuro)science matters for ethics. *Ethics*, 124(4), 695–726. <https://doi.org/10.1086/675875>
- Gromet, D. M., & Darley, J. M. (2009). Punishment and beyond: Achieving justice through the satisfaction of multiple goals. *Law & Society Review*, 43(1), 1–38. <https://doi.org/10.1111/j.1540-5893.2009.00365.x>
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834. <https://doi.org/10.1037/0033-295X.108.4.814>
- Haidt, J., Björklund, F., & Murphy, S. (2000). *Moral dumbfounding: When intuition finds no reason* [Unpublished manuscript].
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Cardenas, J. C., Gurven, M., Gwako, E., Marlowe, F., Tracer, D., Ziker, J., Bolyanatz, A., Henrich, N., & Lesorogol, C. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–1770. <https://doi.org/10.1126/science.1127333>
- Hindriks, F. (2015). How does reasoning (fail to) contribute to moral judgment? Dumbfounding and disengagement. *Ethical Theory and Moral Practice*, 18(2), 237–250. <https://doi.org/10.1007/s10677-015-9575-7>
- Hoffman, M. B., & Goldsmith, T. H. (2004). The biological roots of punishment. *Ohio State Journal of Criminal Law*, 1(2), 627–642.
- Horstmann, N., Hausmann, D., & Ryf, S. (2010). Methods for inducing intuitive and deliberate processing modes. In A. Glöckner & C. Wittmann (Eds.), *Foundations for tracing intuition: Challenges and methods* (pp. 219–237). Psychology Press.
- Kahane, G. (2012). On the wrong track: Process and content in moral psychology: process and content in moral psychology. *Mind & Language*, 27(5), 519–545. <https://doi.org/10.1111/mila.12001>
- Kant, I. (1998). *Grundlegung zur Metaphysik der Sitten* (F.-P. Hansen, Ed.). directmedia. (Original work published 1785).
- Keller, L. B., Oswald, M. E., Stucki, I., & Gollwitzer, M. (2010). A closer look at an eye for an eye: Laypersons' punishment decisions are primarily driven by retributive motives. *Social Justice Research*, 23(2–3), 99–116. <https://doi.org/10.1007/s11211-010-0113-4>
- Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, 118(1), 97–109. <https://doi.org/10.1037/a0020762>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in Linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, Article 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Malle, B. F. (2021). Moral judgments. *Annual Review of Psychology*, 72(1), 293–318. <https://doi.org/10.1146/annurev-psych-072220-104358>
- McFatter, R. M. (1982). Purposes of punishment: Effects of utilities of criminal sanctions on perceived appropriateness. *Journal of Applied Psychology*, 67(3), 255–267. <https://doi.org/10.1037/0021-9010.67.3.255>
- Nadelhoffer, T., Heshmati, S., Kaplan, D., & Nichols, S. (2013). Folk retributivism and the communication confound. *Economics and Philosophy*, 29(2), 235–261. <https://doi.org/10.1017/S0266267113000217>
- Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45(4), 867–872. <https://doi.org/10.1016/j.jesp.2009.03.009>
- Oswald, M. E., & Stucki, I. (2009). A two-process model of punishment. In M. E. Oswald, S. Bieneck, & J. Hupfeld-Heinemann (Eds.), *Social psychology of punishment of crime* (pp. 173–192). Wiley.
- Patil, I., Zucchelli, M. M., Kool, W., Campbell, S., Fornasier, F., Calò, M., Silani, G., Cikara, M., & Cushman, F. (2021). Reasoning supports utilitarian resolutions to moral dilemmas across diverse measures. *Journal of Personality and Social Psychology*, 120(2), 443–460. <https://doi.org/10.1037/pspp0000281>
- Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, 36(1), 163–177. <https://doi.org/10.1111/j.1551-6709.2011.01210.x>
- Peczenik, A. (2005). *Scientia juris: Legal doctrine as knowledge of law and as a source of law*. Springer.
- Petty, R. E., Briñol, P., Loersch, C., & McCaslin, M. J. (2009). The need for cognition. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 318–329). Guilford Press.
- Pizarro, D. A., & Bloom, P. (2003). The intelligence of the moral intuitions: A comment on Haidt (2001). *Psychological Review*, 110(1), 193–196. <https://doi.org/10.1037/0033-295X.110.1.193>
- Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, 9(1), 29–43. <https://doi.org/10.1080/13869790500492466>
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing
- Rehren, P., & Zisman, V. (2021). *Supplemental materials and code for "Testing the intuitive retributivism hypothesis"*. <https://doi.org/10.23668/psycharchives.5009>
- Robinson, P. H., & Darley, J. M. (2007). Intuitions of Justice: Implications for Criminal Law and Justice Policy. *Southern California Law Review*, 81(1), 1–69.
- Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, 10(4), 296–313.
- Sargent, M. J. (2004). Less thought, more punishment: Need for cognition predicts support for punitive responses to crime. *Personality and Social Psychology Bulletin*, 30(11), 1485–1493. <https://doi.org/10.1177/0146167204264481>

- Sauer, H. (2018). *Moral thinking, fast and slow*. Routledge. <https://doi.org/10.4324/9781315467498>
- Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110, Article 104038. <https://doi.org/10.1016/j.jml.2019.104038>
- Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, 141(3), 423–428. <https://doi.org/10.1037/a0025391>
- Stanley, M. L., Yin, S., & Sinnott-Armstrong, W. (2019). A reason-based explanation for moral dumbfounding. *Judgment and Decision Making*, 14(2), 120–129.
- Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55–88). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199230167.003.0003>
- Van Doorn, J., & Brouwers, L. (2017). Third-party responses to injustice: A review on the preference for compensation. *Crime Psychology Review*, 3(1), 59–77. <https://doi.org/10.1080/23744006.2018.1470765>
- Van Knippenberg, A., Dijksterhuis, A., & Vermeulen, D. (1999). Judgement and memory of a criminal act: The effects of stereotypes and cognitive load. *European Journal of Social Psychology*, 29(2–3), 191–201. [https://doi.org/10.1002/\(SICI\)1099-0992\(199903/05\)29:2<191::AID-EJSP923>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1099-0992(199903/05)29:2<191::AID-EJSP923>3.0.CO;2-O)
- Van Prooijen, J.-W. (2010). Retributive versus compensatory justice: Observers' preference for punishing in response to criminal offenses. *European Journal of Social Psychology*, 40, 72–85. <https://doi.org/10.1002/ejsp.611>
- Wood, D. (2010). Punishment: Consequentialism. *Philosophy Compass*, 5(6), 455–469. <https://doi.org/10.1111/j.1747-9991.2010.00287.x>

History

Received April 15, 2020

Revision received April 24, 2021

Accepted April 26, 2021

Published online February 25, 2022

Acknowledgments

This study was preregistered (<https://dx.doi.org/10.23668/psycharchives.3479>). We deviated from our preregistration in our data collection and analysis. We explain and justify these deviations transparently in our methods and analysis sections. The project reported in this paper was presented at a meeting of MAD Lab at Duke University. We thank all participants for their helpful feedback. We also thank Mathias Twardawski and two anonymous reviewers for their helpful comments and suggestions. Finally, we thank the ZPID's PsychLab online for their help with data collection.

Conflict of Interest

The authors declare no conflict of interest.

Open Data

Our study materials and raw data (including codebook, analysis code, and additional analyses/discussion) are available as Electronic Supplementary Material (ESM) at PsychArchives at <https://dx.doi.org/10.23668/psycharchives.5009> (Rehren & Zisman, 2021).

Funding

Open access publication was enabled by the European Research Council (ERC) project “The Enemy of the Good. Towards a Theory of Moral Progress” (grant number: 851043).

ORCID

Paul Rehren

 <https://orcid.org/0000-0003-4267-3097>

Paul Rehren

Ethics Institute

Department of Philosophy and Religious Studies

Utrecht University

Janskerkhof 13

3512 BL Utrecht

The Netherlands

p.h.p.rehren@uu.nl